



PCT/FR 2004 / 002602

REC'D 28 DEC 2004

WIPO

PCT

BREVET D'INVENTION

CERTIFICAT D'UTILITÉ - CERTIFICAT D'ADDITION

COPIE OFFICIELLE

Le Directeur général de l'Institut national de la propriété industrielle certifie que le document ci-annexé est la copie certifiée conforme d'une demande de titre de propriété industrielle déposée à l'Institut.

Fait à Paris, le 19 OCT. 2004

Pour le Directeur général de l'Institut
national de la propriété industrielle
Le Chef du Département des brevets

Martine PLANCHE

DOCUMENT DE PRIORITÉ

**PRÉSENTÉ OU TRANSMIS
CONFORMÉMENT À LA
RÈGLE 17.1.a) OU b)**

**INSTITUT
NATIONAL DE
LA PROPRIÉTÉ
INDUSTRIELLE**

SIEGE

26 bis, rue de Saint-Petersbourg
75800 PARIS cedex 08
Téléphone : 33 (0)1 53 04 53 04
Télécopie : 33 (0)1 53 04 45 23
www.inpi.fr



INSTITUT NATIONAL DE LA PROPRIÉTÉ INDUSTRIELLE
15, rue de Saint Pétersbourg
90 Paris Cedex 08
phone : 01 53 04 53 04 Télécopie : 01 42 94 86 54

BREVET D'INVENTION CERTIFICAT D'UTILITÉ

Code de la propriété intellectuelle - Livre VI

cerfa
N° 11354*01

REQUÊTE EN DÉLIVRANCE 1/2

Cet imprimé est à remplir lisiblement à l'encre noire

08 540 W / 260899

Réservé à l'INPI

ÉMISSION DES PIÈCES
DE

EU

° D'ENREGISTREMENT

ATIONAL ATTRIBUÉ PAR L'INPI

ATE DE DÉPÔT ATTRIBUÉE

AR L'INPI

los références pour ce dossier
facultatif) IFB 03 CA CNR GLOB

NOM ET ADRESSE DU DEMANDEUR OU DU MANDATAIRE
À QUI LA CORRESPONDANCE DOIT ÊTRE ADRESSÉE

GROSSET-FOURNIER & DEMACHY
54, rue Saint-Lazare
F-75009 Paris

Confirmation d'un dépôt par télécopie

☐ N° attribué par l'INPI à la télécopie

2 NATURE DE LA DEMANDE

Cochez l'une des 4 cases suivantes

Demande de brevet

☒

Demande de certificat d'utilité

☐

Demande divisionnaire

☐

*Demande de brevet initiale
ou demande de certificat d'utilité initiale*

N°

Date | / /

N°

Date | / /

Transformation d'une demande de

brevet européen *Demande de brevet initiale*

☐

N°

Date | / /

3 TITRE DE L'INVENTION (200 caractères ou espaces maximum)
PROCÉDE DE DISSOCIATION DE LA MOLECULE D'HEMOGLOBINE EXTRACELLULAIRE D'ARENICOLA MARINA, CHAINES
PROTEIQUES CONSTITUANT LADITE MOLECULE ET SEQUENCES NUCLEOTIDIQUES CODANT POUR LESDITES CHAINES
PROTEIQUES

4 DÉCLARATION DE PRIORITÉ

OU REQUÊTE DU BÉNÉFICE DE

LA DATE DE DÉPÔT D'UNE

DEMANDE ANTÉRIEURE FRANÇAISE

Pays ou organisation

Date | / /

N°

Pays ou organisation

Date | / /

N°

Pays ou organisation

Date | / /

N°

☐ S'il y a d'autres priorités, cochez la case et utilisez l'imprimé «Suite»

5 DEMANDEUR

☐ S'il y a d'autres demandeurs, cochez la case et utilisez l'imprimé «Suite»

Nom ou dénomination sociale

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

Prénoms

Forme juridique

N° SIREN

Code APE-NAF

Adresse

Rue

Code postal et ville

3, rue Michel-Ange

F-75794 PARIS CEDEX 16

Pays

FRANCE

Nationalité

FRANCAISE

N° de téléphone (facultatif)

N° de télécopie (facultatif)

Adresse électronique (facultatif)

BREVET D'INVENTION
CERTIFICAT D'UTILITÉ

 INSTITUT
 NATIONAL DE
 LA PROPRIÉTÉ
 INDUSTRIELLE

 4 OCT 2003
 75 INPI PARIS

REQUÊTE EN DÉLIVRANCE 2/2

0311992

Réservé à l'INPI

 REMISE DES PIÈCES
 DATE

LIEU

 N° D'ENREGISTREMENT
 NATIONAL ATTRIBUÉ PAR L'INPI

DB 540 W / 250899

 Vos références pour ce dossier :
 (facultatif)

IFB 03 CA CNR GLOB

6 MANDATAIRE

Nom

GROSSET-FOURNIER

Prénom

Chantal

Cabinet ou Société

GROSSET-FOURNIER & DEMACHY

 N° de pouvoir permanent et/ou
 de lien contractuel

Adresse

Rue

54, rue Saint-Lazare

Code postal et ville

75009 PARIS

N° de téléphone (facultatif)

01.42.81.09.58

N° de télécopie (facultatif)

01.42.81.08.71

Adresse électronique (facultatif)

7 INVENTEUR (S)

Les inventeurs sont les demandeurs

☐ Oui☒ Non Dans ce cas fournir une désignation d'inventeur(s) séparée**8 RAPPORT DE RECHERCHE**

Uniquement pour une demande de brevet (y compris division et transformation)

 Établissement immédiat
 ou établissement différé
☒☐

Paiement échelonné de la redevance

Paiement en deux versements, uniquement pour les personnes physiques

☐ Oui☐ Non**9 RÉDUCTION DU TAUX
DES REDEVANCES**

Uniquement pour les personnes physiques

☐ Requête pour la première fois pour cette invention (joindre un avis de non-imposition)☐ Requête antérieurement à ce dépôt (joindre une copie de la décision d'admission pour cette invention ou indiquer sa référence):
 Si vous avez utilisé l'imprimé «Suite»,
 indiquez le nombre de pages jointes
**10 SIGNATURE DU DEMANDEUR
OU DU MANDATAIRE
(Nom et qualité du signataire)**
 Chantal, Catherine GROSSET-FOURNIER
 Mandataire
 422.5/PP.112

 VISA DE LA PRÉFECTURE
 OU DE L'INPI

PROCÉDÉ DE DISSOCIATION DE LA MOLÉCULE D'HÉMOGLOBINE EXTRACELLULAIRE D'*ARENICOLA MARINA*, CHAÎNES PROTÉIQUES CONSTITUANT LADITE MOLÉCULE ET SÉQUENCES NUCLÉOTIDIQUES CODANT POUR LESDITES CHAÎNES PROTÉIQUES

La présente invention a pour objet un procédé de dissociation de la molécule d'hémoglobine extracellulaire d'Annélides, notamment d'*Arenicola marina*, ainsi que les chaînes protéiques constituant ladite molécule.

La présente invention a également pour objet les séquences nucléotidiques codant pour les chaînes protéiques susmentionnées, ainsi que le procédé de préparation de ces séquences nucléotidiques.

Le sang est un liquide complexe dont la fonction principale est de transporter l'oxygène et le gaz carbonique, afin d'assurer les processus respiratoires. C'est la molécule d'hémoglobine, que l'on trouve dans les hématies, qui assure cette fonction.

La molécule d'hémoglobine de mammifère est formée par l'assemblage de quatre chaînes polypeptidiques fonctionnelles semblables deux à deux (2 chaînes de globine de type α et 2 chaînes de globine de type β). Chacune de ces chaînes polypeptidiques possède la même structure tertiaire d'une molécule de myoglobine (Dickerson et Geis, 1983).

L'hème, site actif de l'hémoglobine, est un anneau protoporphyrinique tétrapyrrolique, contenant en son centre un unique atome ferreux. L'atome de fer, fixateur de l'oxygène, contracte 6 liaisons de coordinence : quatre avec les atomes d'azote de la porphyrine, une avec l'histidine proximale F8₁ et une avec la molécule d'oxygène lors de l'oxygénation de la globine.

On est actuellement confronté aux problèmes d'approvisionnement de sang, dus à la diminution des dons du sang par peur de contamination. Ainsi, la recherche de substituts sanguins s'est accélérée au cours des dernières années. On cherche à concevoir des substituts sanguins artificiels capables d'éliminer les risques de transmission de maladies infectieuses, mais également à s'affranchir des problèmes de compatibilité des groupes sanguins.

Jusqu'à présent, les principales voies de recherche concernent la synthèse de produits chimiques d'une part (Clark et Gollan, 1966) et la synthèse de produits biologiques d'autre part (Chang, 1957 ; Chang, 1964).

En ce qui concerne la première voie de recherche, on a utilisé les perfluorocarbones (PFC). Les PFC sont des produits chimiques capables de transporter l'oxygène et ils peuvent dissoudre une grande quantité de gaz, comme l'oxygène et le dioxyde de carbone.

5 Actuellement, on essaye de produire des émulsions de ces produits qui pourraient être dispersés dans le sang d'une façon plus efficace (Reiss, 1991 ; Reiss, 1994 ; Goodin et al., 1994).

10 L'avantage des PFC demeure dans leur capacité oxyphorique directement proportionnelle à la quantité d'oxygène se trouvant dans les poumons. Par ailleurs, en raison de l'absence de membrane à traverser, les PFC peuvent transporter l'oxygène plus rapidement vers les tissus. Toutefois, on ne connaît pas les effets à long terme de la rétention de ces produits dans l'organisme. Lorsque ces produits ont été utilisés pour la première fois comme substitut sanguin dans les années 1960 chez la souris (Clark et Gollan, 1966 ; Geyer et al., 1966 ; Sloviter et Kamimoto, 1967), les effets secondaires
15 ont été très importants. Les PFC n'étaient pas éliminés de la circulation d'une façon satisfaisante et s'accumulaient dans les tissus de l'organisme, ce qui provoquait des œdèmes.

Dans les années 1980, on a testé une nouvelle version de PFC en phase clinique. Mais les problèmes de stockage, de coût financier, d'effets secondaires non négligeables
20 et la faible efficacité de ce composé ont empêché l'extension de sa commercialisation (Naito, 1978 ; Mitsuno et Naito, 1979 ; Mitsuno et Ohyanagi, 1985).

Récemment, on a mis au point une nouvelle génération de PFC (PFBO perfluorooctylbromide). Un nouveau produit (Reiss, 1991) est en cours de test clinique aux Etats-Unis, mais on a déjà constaté qu'une augmentation de la quantité d'oxygène
25 dans le sang peut engendrer une accumulation d'oxygène dans les tissus, ce qui est dangereux pour l'organisme (formation d'oxygène radicalaire de type superoxyde).

Ainsi, malgré les progrès réalisés au fur et à mesure, les effets secondaires de ces composés sont encore trop importants pour être commercialisés à grande échelle.

En ce qui concerne la deuxième voie de recherche, des chercheurs ont travaillé sur
30 la mise au point de substituts sanguins en modifiant la structure de l'hémoglobine naturelle (Chang, 1957 ; Chang, 1997). Pour obtenir un substitut sanguin de type hémoglobine modifiée, on utilise des hémoglobines provenant de microorganismes génétiquement modifiés, ou d'origine humaine ou animale, notamment la molécule d'hémoglobine bovine. En effet, l'hémoglobine bovine est légèrement différente de

l'hémoglobine humaine sur le plan immunologique, mais elle transporte plus facilement l'oxygène vers les tissus. Néanmoins, le risque de contamination virale ou de type encéphalopathie spongiforme reste toujours important.

Pour être fonctionnelle, l'hémoglobine doit être en contact avec un effecteur allostérique le 2,3-diphosphoglycérate (2,3-DPG), présent uniquement à l'intérieur des globules rouges (Dickerson et Geis, 1983). De plus, sans le 2,3-DPG et d'autres éléments présents dans les globules rouges comme la méthémoglobine réductase, l'hémoglobine subit un processus d'auto-oxydation et perd sa capacité à transporter de l'oxygène ou du dioxyde de carbone.

On peut éliminer ces processus en modifiant la structure de l'hémoglobine, et plus précisément en stabilisant les liaisons faibles de la molécule tétramérique entre les deux dimères α et β (Chang, 1971). De nombreuses modifications ont été testées : liaison covalente entre deux chaînes α , entre deux chaînes β ou encore entre α et β (Payne, 1973 ; Bunn et Jandl, 1968).

On a également tenté de polymériser les molécules tétramériques ou de les conjuguer avec un polymère nommé polyéthylène glycol (PEG) (Nho et al., 1992). Ces modifications ont pour conséquence de stabiliser la molécule et d'augmenter sa taille, empêchant son élimination par les reins.

On a beaucoup étudié les Annélides pour leur hémoglobine extracellulaire (Terwilliger, 1992 ; Lamy et al., 1996). Ces molécules d'hémoglobine extracellulaire sont présentes chez les trois classes d'Annélides : Polychètes, Oligochètes et Achètes et même chez les Vestimentifères. Ce sont des biopolymères géants, constitués d'environ 200 chaînes polypeptidiques appartenant à 6 ou 7 types différents que l'on regroupe généralement en deux catégories. La première catégorie, comptant 144 à 192 éléments, regroupe les chaînes polypeptidiques dites "fonctionnelles" portant un site actif et capables de lier réversiblement l'oxygène ; ce sont des chaînes de type globine dont les masses sont comprises entre 15 et 18 kDa et qui sont très similaires aux chaînes de type α et β de vertébrés. La deuxième catégorie, comptant 36 à 42 éléments, regroupe les chaînes polypeptidiques dites de "structure" possédant peu ou pas de site actif mais permettant l'assemblage des douzièmes.

Les premières images obtenues sur des hémoglobines extracellulaires d'Arénicole (Roche et al., 1960) ont révélé des éléments hexagonaux. Chaque molécule d'hémoglobine est constituée de deux hexagones superposés (Levin, 1963 ; Roche,

1965) que l'on a nommés bicouche hexagonale (hexagonal bilayer) et chaque hexagone est lui-même formé par l'assemblage de six éléments en forme de goutte d'eau (Van Bruggen et Weber, 1974 ; Kapp et Crewe, 1984); nommés structure globulaire creuse (hollow globular structure)(De Haas et al., 1996) ou "douzième". La molécule native est formée de douze de ces sous-unités, d'une masse moléculaire d'environ 250 kDa.

Ainsi, le brevet français n°2 809 624 concerne l'utilisation comme substitut sanguin d'hémoglobine extracellulaire d'*Arenicola marina*, un Annélide Polychète de l'écosystème intertidal, ledit substitut sanguin permettant d'éliminer les problèmes de pénurie de dons.

Bien que l'architecture globale de l'hémoglobine de *Arenicola marina* soit connue, notamment grâce à son étude quaternaire fine par spectrométrie de masse (Zal et al., 1997), les séquences primaires des différentes chaînes protéiques qui la composent ne le sont pas.

Ainsi, la présente invention a pour but de fournir les séquences protéiques qui composent la molécule d'hémoglobine d'*Arenicola marina*.

Un autre but de la présente invention consiste à fournir les premières étapes de la synthèse *in vitro* d'hémoglobine extracellulaire d'*Arenicola marina* afin d'élaborer un substitut sanguin au moyen de procédé de biochimie et de biologie moléculaire.

Un autre but de la présente invention consiste à fournir un procédé de préparation de la molécule d'hémoglobine, éventuellement simplifiée, par génie génétique, et ce afin notamment d'augmenter le stock de cette molécule dans le cadre d'une utilisation comme substitut sanguin.

La présente invention concerne un procédé de dissociation de la molécule d'hémoglobine extracellulaire d'Annélides, notamment d'*Arenicola marina*, permettant l'obtention des dix chaînes protéiques constituant ladite molécule,

ledit procédé étant caractérisé en ce qu'il comprend une étape de mise en présence d'un échantillon d'hémoglobine extracellulaire d'Annélides, notamment d'*Arenicola marina* avec au moins un agent dissociant, notamment un mélange constitué de dithiothréitol (DTT) ou de chlorhydrate de tris(2-carboxyéthyl)phosphine (TCEP) ou de bêta-mercaptoéthanol et d'un tampon de dissociation, pendant un temps suffisant pour séparer les chaînes protéiques les unes des autres.

Par "hémoglobine extracellulaire", on désigne une hémoglobine non contenue dans les cellules et dissoute dans le sang.

L'expression "agent dissociant" désigne un composé chimique capable de rompre les interactions faibles (hydrophobes, électrostatiques, hydrogènes...) ou fortes (ponts disulfures).

Par "tampon de dissociation", on désigne un liquide contenant un tampon permettant d'ajuster le pH et contenant des agents dissociants.

Les dix chaînes protéiques constituant la molécule d'hémoglobine extracellulaire d'*Arenicola marina* comprennent 8 chaînes de type globine et 2 chaînes de type structure.

On rappelle que l'hémoglobine extracellulaire de *Arenicola marina* d'une masse de 3648 ± 24 kDa est constituée de 198 chaînes polypeptidiques appartenant à 10 types différents regroupés en deux catégories :

- la première (156 chaînes) regroupe des chaînes polypeptidiques dites "fonctionnelles" portant un site actif capable de lier réversiblement l'oxygène ; ce sont des chaînes de type globine dont les masses sont comprises entre 15 et 18 kDa ; ces chaînes sont très similaires aux chaînes de types α et β des vertébrés ; et

- la deuxième (42 chaînes) regroupe des chaînes polypeptidiques dites de "structure" (ou "linker") possédant peu ou pas de site actif mais permettant l'assemblage des dodécamères ; ces chaînes ont des masses moléculaires comprises entre 22 et 27 kDa.

La présente invention concerne un procédé de dissociation de la molécule d'hémoglobine extracellulaire d'*Arenicola marina*, permettant l'obtention des dix chaînes protéiques constituant ladite molécule,

ledit procédé étant caractérisé en ce qu'il comprend une étape de mise en présence d'un échantillon d'hémoglobine extracellulaire d'*Arenicola marina* avec un mélange constitué de dithiothréitol (DTT) et d'un tampon de dissociation, pendant environ une heure à trois semaines.

Un procédé de dissociation avantageux de l'invention est caractérisé en ce que le tampon de dissociation comprend 0,1 M de base Trisma (tris[hydroxyméthyl]aminométhane) et 0 à 10 mM d'EDTA ajusté à un pH compris d'environ 5 à environ 12.

De préférence, ledit tampon de dissociation comprend de l'EDTA à une concentration d'environ 1 mM ajusté à un pH d'environ 10 avec une solution de soude 2N.

Selon un mode de réalisation avantageux, le procédé de dissociation de l'invention est caractérisé en ce que les dix chaînes protéiques constituant ladite

molécule sont obtenues par la dissociation de quatre sous-unités par exemple en présence de DTT, lesdites sous-unités étant elles-mêmes obtenues par mise en présence d'un échantillon d'hémoglobine extracellulaire d'*Arenicola marina* avec différents agents dissociants, notamment un tampon de dissociation.

La molécule native se dissocie en sous-unités sous l'action d'agents dissociants non réducteurs. Il n'y a donc pas rupture des liaisons covalentes. Toutefois, après l'action d'un agent réducteur (coupure des liaisons covalentes), les sous-unités sont réduites en chaînes polypeptidiques (chaînes protéiques constituées de l'assemblage d'acides aminés).

Les 4 sous-unités susmentionnées sont donc : les monomères, les dimères, les trimères et les dodécamères.

Les monomères sont des chaînes de globine.

Les dimères sous forme homo et hétérodimères sont des chaînes de structure.

Les trimères sont des assemblages covalents de trois chaînes de globine.

Les dodécamères sont constitués de 12 chaînes protéiques ; par exemple : 3 trimères + 3 monomères, 2 trimères + 6 monomères, 1 trimère + 9 monomères.

On peut donc obtenir les 10 chaînes protéiques soit en une seule étape par dissociation directe de l'hémoglobine extracellulaire d'*Arenicola marina*, soit en deux étapes, l'une consistant en la dissociation de l'hémoglobine extracellulaire d'*Arenicola marina* en 4 sous-unités et l'autre en la dissociation desdites 4 sous-unités en dix chaînes protéiques.

La présente invention concerne également un procédé de dissociation tel que défini ci-dessus, caractérisé en ce que les agents dissociants utilisés pour obtenir les 4 sous-unités susmentionnées sont les suivants :

- une solution tampon de dissociation comprenant : 0,1 M de base Trisma (tris[hydroxyméthyl]aminométhane) et 0 à 10 mM d'EDTA ajusté à un pH compris d'environ 5 à environ 12, et
- une solution d'urée dont la concentration est comprise d'environ 0,1 M à environ 8 M, et est notamment égale à 3 M.

La présente invention concerne également un procédé de dissociation tel que défini ci-dessus, caractérisé en ce que les agents dissociants pour obtenir les 4 sous-unités sont les suivants :

- une solution tampon de dissociation comprenant 0,1 M de base Trisma et 1 mM d'EDTA ajusté à pH 10, et

– 3 M d'urée.

Le procédé de dissociation de l'invention permet d'obtenir une composition contenant le mélange des dix chaînes protéiques constituant la molécule d'hémoglobine extracellulaire d'*Arenicola marina*.

La présente invention concerne également un procédé de préparation de couples d'amorces à partir des chaînes protéiques telles qu'obtenues selon le procédé tel que défini ci-dessus, ledit procédé étant caractérisé en ce qu'il comprend les étapes suivantes :

– l'isolement de chacune des dix chaînes protéiques constituant la molécule d'hémoglobine telles qu'obtenues selon le procédé tel que défini ci-dessus,

– le microséquençage de chacune des dix chaînes protéiques isolées susmentionnées par spectrométrie de masse, afin d'obtenir une microséquence correspondant à chacune des dix séquences constituées de 5 à 20 acides aminés, et

– la détermination de dix couples d'amorces dégénérées à partir des microséquences susmentionnées.

La première étape d'isolement des chaînes protéiques est notamment effectuée par chromatographie liquide d'exclusion, anionique et phase inverse à partir du mélange susmentionné comprenant les dix chaînes protéiques constituant la molécule d'hémoglobine telles qu'obtenues selon le procédé de dissociation de l'invention.

L'expression "microséquence" désigne des fragments de séquences protéiques.

Les microséquences susmentionnées peuvent provenir à la fois des extrémités C- et N-terminales mais aussi de séquences internes.

Ces microséquences peuvent être obtenues par chromatographie liquide en phase inverse ou à partir de gel 2D à partir d'hémoglobine purifiée d'*Arenicola marina*. Chaque pic ou tache a été découpé et digéré par une protéase. Les peptides ainsi obtenus ont été extraits des gels et séparés par capLC (chromatographie liquide capillaire). Les fragments sont ensuite analysés par spectrométrie de masse.

L'expression "couples d'amorces dégénérées" désigne des séquences nucléotidiques obtenues à partir de fragments de séquences protéiques. On parle d'amorces dégénérées en raison de la dégénérescence du code génétique (plusieurs codons pour 1 acide aminé).

La dernière étape de détermination des couples d'amorces dégénérées correspond à leur synthèse.

Cette étape permet d'obtenir à la fois des amorces sens et des amorces antisens.

La présente invention concerne également des couples d'amorces tels qu'obtenus selon le procédé tel que défini ci-dessus, lesdits couples étant notamment les suivants :

a) *Amorce sens* : GAR TGY GGN CCN TTR CAR CG SEQ ID NO : 9

Amorce antisens : CCA NGC NTC YTT RTC RAA GCA SEQ ID NO : 10

b) *Amorce sens* : AN TGY GGN CCN CTN CAR CG SEQ ID NO : 11

Amorce antisens : CCA NGC NTC YTT RTC RAA GCA SEQ ID NO : 12

c) *Amorce sens* : AAR GTI AAR CAN AAC TGG SEQ ID NO : 13

Amorce antisens : CCA NGC NCC DAT RTC RAA SEQ ID NO : 14

d) *Amorce sens* : TGY TGY AGY ATH GAR GAY CG SEQ ID NO : 15

Amorce antisens : CA NGC NYC RCT RTT RAA RCA SEQ ID NO : 16

e) *Amorce sens* : GAY GAY TGY GTN TGY CC SEQ ID NO : 17

Amorce antisens : ANA CRA ART CRT CNC CRC A SEQ ID NO : 18

où :

R représente A ou G,

Y représente C ou T,

N représente A, G, C ou T,

I représente l'inosine,

D représente A, G ou T,

H représente A, C ou T.

La présente invention concerne également un procédé de préparation de séquences nucléotidiques codant pour les chaînes protéiques constituant la molécule d'hémoglobine d'*Arenicola marina*, à partir des amorces telles qu'obtenues selon le procédé tel que défini ci-dessus, ledit procédé étant caractérisé en ce qu'il correspond à un procédé d'amplification en chaîne par polymérase (PCR), comprenant une répétition d'au moins 30 fois du cycle constitué par les étapes suivantes :

* la dénaturation, par chauffage, de l'ADNc codant pour chacune des dix chaînes protéiques constituant la molécule d'hémoglobine d'*Arenicola marina*, de façon à séparer l'ADNc en deux brins monocaténares, ledit ADNc étant obtenu à partir d'ARNm mature,

* l'hybridation des couples d'amorces tels qu'obtenus par le procédé tel que défini ci-dessus aux brins d'ADNc monocaténares susmentionnés à une température adéquate, pour obtenir des amorces hybridées, et

* l'élongation des amorces hybridées par une polymérase à une température adéquate.

L'ADNc codant pour les dix chaînes protéiques susmentionnées est obtenu à partir d'ARNm mature, ledit ARNm mature étant obtenu par purification à partir d'ARN totaux extraits sur des Arénicoles juvéniles en pleine croissance, lesdits Arénicoles juvéniles présentant un taux élevé d'expression des différents ARN messagers. Tout organisme en pleine croissance va synthétiser plus de protéines donc plus d'ARNm.

Si le cycle susmentionné est répété moins de 30 fois, l'amplification de l'ADN est réduite.

L'expression "dénaturation par chauffage" désigne la déstructuration de la double hélice d'ADN en deux brins monocaténares.

L'expression "hybridation à une température adéquate" désigne la reconnaissance par les amorces de leurs séquences complémentaires sur l'ADN matrice.

L'expression "élongation à une température adéquate" désigne la synthèse de brin d'ADNc par la polymérase.

Selon un mode de réalisation avantageux, le procédé de préparation de l'invention est caractérisé en ce que :

- la première étape dudit procédé est une étape de dénaturation de l'ADNc codant pour chacune des dix chaînes protéiques constituant la molécule d'hémoglobine d'*Arenicola marina* d'environ 10 secondes à environ 5 minutes à une température comprise d'environ 90°C à environ 110°C,

- le cycle, répété environ 30 à 40 fois, comprend les étapes suivantes :

* une étape de dénaturation de l'ADNc codant pour chacune des dix chaînes protéiques constituant la molécule d'hémoglobine d'*Arenicola marina* d'environ 10 secondes à environ 5 minutes, à une température comprise d'environ 90°C à environ 110°C,

* une étape d'hybridation des couples d'amorces de l'invention aux brins d'ADNc monocaténares susmentionnés pour obtenir des amorces

hybridées, d'environ 20 secondes à environ 2 minutes, à une température comprise d'environ 50°C à environ 56°C,

* une étape d'élongation des amorces hybridées telles qu'obtenues précédemment par une polymérase d'environ 20 secondes à environ 1 minute et 30 secondes, à une température comprise d'environ 70°C à environ 75°C, et

— la dernière étape du procédé est une étape d'élongation des amorces hybridées telles qu'obtenues précédemment par une polymérase d'environ 5 minutes à environ 15 minutes à une température comprise d'environ 70°C à environ 75°C.

La réaction de PCR du procédé de l'invention est notamment effectuée en présence d'ADNc (1 µl), d'amorce sens (100 ng) et antisens (100 ng), de dNTP (200 µM), de MgCl₂ (2 mM), de tampon PCR (fourni avec la polymérase) (1 X, à la concentration du tampon), de la Taq polymérase (1 unité) et d'eau (25 µl qsp).

La présente invention concerne également un procédé de préparation tel que défini ci-dessus, caractérisé en ce que les couples d'amorces utilisés sont tels que définis précédemment.

Un procédé de préparation particulièrement avantageux selon l'invention est un procédé de préparation tel que défini ci-dessus, caractérisé en ce que le couple d'amorces utilisé est le suivant : (GAR TGY GGN CCN TTR CAR CG ; CCA NGC NTC YTT RTC RAA GCA), R, Y et N étant tels que définis ci-dessus, ledit procédé étant caractérisé en ce que :

— la première étape du procédé est une étape de dénaturation de l'ADNc codant pour chacune des dix chaînes protéiques constituant la molécule d'hémoglobine d'*Arenicola marina*, de 4 minutes à une température égale à 95°C,

— le cycle, répété 35 fois, comprend les étapes suivantes :

* une étape de dénaturation de l'ADNc codant pour chacune des dix chaînes protéiques constituant la molécule d'hémoglobine d'*Arenicola marina*, de 30 secondes à une température égale à 95°C,

* une étape d'hybridation des couples d'amorces de l'invention aux brins d'ADNc monocaténares susmentionnés pour obtenir des amorces hybridées, de 30 secondes à une température égale à 56°C.

* une étape d'élongation des amorces hybridées telles qu'obtenues précédemment par une polymérase de 40 secondes à une température égale à 72°C, et

– la dernière étape du procédé est une étape d'élongation des amorces hybridées telles qu'obtenues précédemment par une polymérase de 10 minutes à une température égale à 72°C,

afin d'obtenir la séquence nucléotidique SEQ ID NO : 1.

La séquence nucléotidique SEQ ID NO : 1 est une nouvelle séquence nucléotidique codant pour une chaîne protéique correspondant à une chaîne de globine, notée A2 (a). SEQ ID NO : 1 comprend 376 nucléotides.

Un procédé de préparation particulièrement avantageux selon l'invention est un procédé de préparation tel que défini ci-dessus, caractérisé en ce que le couple d'amorces utilisé est le suivant : (AN TGY GGN CCN CTN CAR CG ; CCA NGC NTC YTT RTC RAA GCA), N, Y et R étant tels que définis ci-dessus,

ledit procédé étant caractérisé en ce que :

– la première étape du procédé est une étape de dénaturation de l'ADNc codant pour chacune des dix chaînes protéiques constituant la molécule d'hémoglobine d'*Arenicola marina*, de 4 minutes à une température égale à 95°C,

– le cycle, répété 35 fois, comprend les étapes suivantes :

* une étape de dénaturation de l'ADNc codant pour chacune des dix chaînes protéiques constituant la molécule d'hémoglobine d'*Arenicola marina*, de 30 secondes à une température égale à 95°C,

* une étape d'hybridation des couples d'amorces de l'invention aux brins d'ADNc monocaténares susmentionnés pour obtenir des amorces hybridées, de 30 secondes à une température égale à 52°C,

* une étape d'élongation des amorces hybridées telles qu'obtenues précédemment par une polymérase de 40 secondes à une température égale à 72°C, et

– la dernière étape du procédé est une étape d'élongation des amorces hybridées telles qu'obtenues précédemment par une polymérase de 10 minutes à une température égale à 72°C,

afin d'obtenir la séquence nucléotidique SEQ ID NO : 3.

La séquence nucléotidique SEQ ID NO : 3 est une nouvelle séquence nucléotidique codant pour une chaîne protéique correspondant à une chaîne de globine, notée A2 (b). SEQ ID NO : 3 comprend 288 nucléotides.

Un procédé de préparation particulièrement avantageux selon l'invention est un procédé de préparation tel que défini ci-dessus, caractérisé en ce que le couple d'amorces utilisé est le suivant : (AAR GTI AAR CAN AAC TGG ; CCA NGC NCC DAT RTC RAA), R, I, N et D étant tels que définis ci-dessus,

ledit procédé étant caractérisé en ce que :

- la première étape du procédé est une étape de dénaturation de l'ADNc codant pour chacune des dix chaînes protéiques constituant la molécule d'hémoglobine d'*Arenicola marina*, de 4 minutes à une température égale à 95°C,

- le cycle, répété 35 fois, comprend les étapes suivantes :

- * une étape de dénaturation de l'ADNc codant pour chacune des dix chaînes protéiques constituant la molécule d'hémoglobine d'*Arenicola marina*, de 1 minute à une température égale à 95°C,

- * une étape d'hybridation des couples d'amorces de l'invention aux brins d'ADNc monocaténares susmentionnés pour obtenir des amorces hybridées, de 1 minute à une température égale à 50°C,

- * une étape d'élongation des amorces hybridées telles qu'obtenues précédemment par une polymérase de 1 minute et 30 secondes à une température égale à 72°C, et

- la dernière étape du procédé est une étape d'élongation des amorces hybridées telles qu'obtenues précédemment par une polymérase de 10 minutes à une température égale à 72°C,

afin d'obtenir la séquence nucléotidique SEQ ID NO : 5.

La séquence nucléotidique SEQ ID NO : 5 est une nouvelle séquence nucléotidique codant pour une chaîne protéique correspondant à une chaîne de globine, notée A1. SEQ ID NO : 5 comprend 360 nucléotides.

Un procédé de préparation particulièrement avantageux selon l'invention est un procédé de préparation tel que défini ci-dessus, caractérisé en ce que le couple d'amorces utilisé est le suivant : (TGY TGY AGY ATH GAR GAY CG ; CA NGC NYC RCT RTT RAA RCA), Y, H, R et N étant tels que définis ci-dessus, le dit procédé étant caractérisé en ce que :

- la première étape du procédé est une étape de dénaturation de l'ADNc codant pour chacune des dix chaînes protéiques constituant la molécule d'hémoglobine d'*Arenicola marina*, de 4 minutes à une température égale à 95°C,

- le cycle, répété 35 fois, comprend les étapes suivantes :

* une étape de dénaturation de l'ADNc codant pour chacune des dix chaînes protéiques constituant la molécule d'hémoglobine d'*Arenicola marina*, de 30 secondes à une température égale à 95°C,

* une étape d'hybridation des couples d'amorces de l'invention aux brins d'ADNc monocaténares susmentionnés pour obtenir des amorces hybridées, de 40 secondes à une température égale à 52°C,

* une étape d'élongation des amorces hybridées telles qu'obtenues précédemment par une polymérase de 30 secondes à une température égale à 72°C, et

- la dernière étape du procédé est une étape d'élongation des amorces hybridées telles qu'obtenues précédemment par une polymérase de 10 minutes à une température égale à 72°C,

afin d'obtenir la séquence nucléotidique SEQ ID NO : 7.

La séquence nucléotidique SEQ ID NO : 7 est une nouvelle séquence nucléotidique codant pour une chaîne protéique correspondant à une chaîne de globine, notée B2. SEQ ID NO : 7 comprend 390 nucléotides.

La présente invention concerne également des protéines codées par l'une des séquences telles qu'obtenues selon le procédé tel que défini ci-dessus.

Une protéine préférée selon l'invention est une protéine telle que définie ci-dessus, caractérisée en ce qu'elle comprend ou est constituée par :

- la séquence SEQ ID NO : 2,

- ou toute séquence dérivée de la séquence SEQ ID NO : 2 ou d'un fragment défini ci-dessous, notamment par substitution, suppression ou addition d'un ou plusieurs acides aminés, sous réserve que ladite séquence dérivée permette le transport de l'oxygène,

- ou toute séquence homologue de la séquence SEQ ID NO : 2 ou d'un fragment défini ci-dessous, ayant de préférence une homologie d'au moins environ 75%, notamment d'au moins environ 85%, avec la séquence SEQ ID NO : 2, sous réserve que ladite séquence homologue permette le transport de l'oxygène,

— ou tout fragment d'une des séquences définies ci-dessus, sous réserve que ledit fragment permette le transport de l'oxygène, notamment tout fragment étant constitué d'au moins environ 60 acides aminés, et notamment d'au moins environ 160 acides aminés contigus dans la séquence SEQ ID NO : 2.

5 La séquence SEQ ID NO : 2 est une nouvelle séquence protéique correspondant à une chaîne de globine, notée A2 (a).

— Les propriétés de transport de l'oxygène des séquences protéiques de l'invention peuvent être notamment vérifiées en mesurant leur spectre d'absorption par spectrophotométrie typique de l'oxyhémoglobine.

10 Une protéine préférée selon l'invention est une protéine telle que définie ci-dessus, caractérisée en ce qu'elle comprend ou est constituée par :

— la séquence SEQ ID NO : 4,

— ou toute séquence dérivée de la séquence SEQ ID NO : 4 ou d'un fragment défini ci-dessus, notamment par substitution, suppression ou addition d'un ou plusieurs acides aminés, sous réserve sous réserve que ladite séquence dérivée permette le transport de l'oxygène,

15 — ou toute séquence homologue de la séquence SEQ ID NO : 4 ou d'un fragment défini ci-dessus, ayant de préférence une homologie d'au moins environ 75%, notamment d'au moins environ 85%, avec la séquence SEQ ID NO : 4, sous réserve que ladite séquence homologue permette le transport de l'oxygène,

20 — ou tout fragment d'une des séquences définies ci-dessus, sous réserve que ledit fragment permette le transport de l'oxygène, notamment tout fragment étant constitué d'au moins environ 60 acides aminés, et notamment d'au moins environ 160 acides aminés contigus dans la séquence SEQ ID NO : 4.

25 La séquence SEQ ID NO : 4 est une nouvelle séquence protéique correspondant à une chaîne de globine, notée A2 (b).

Une protéine préférée selon l'invention est une protéine telle que définie ci-dessus, caractérisée en ce qu'elle comprend ou est constituée par :

— la séquence SEQ ID NO : 6,

30 — ou toute séquence dérivée de la séquence SEQ ID NO : 6 ou d'un fragment défini ci-dessus, notamment par substitution, suppression ou addition d'un ou plusieurs acides aminés, sous réserve sous réserve que ladite séquence dérivée permette le transport de l'oxygène,

– ou toute séquence homologue de la séquence SEQ ID NO : 6 ou d'un fragment défini ci-dessous; ayant de préférence une homologie d'au moins environ 75%, notamment d'au moins environ 85%, avec la séquence SEQ ID NO : 6, sous réserve que ladite séquence homologue permette le transport de l'oxygène,

5 – ou tout fragment d'une des séquences définies ci-dessus, sous réserve que ledit fragment permette le transport de l'oxygène, notamment tout fragment étant constitué d'au moins environ 60 acides aminés, et notamment d'au moins environ 160 acides aminés contigus dans la séquence SEQ ID NO : 6.

10 La séquence SEQ ID NO : 6 est une nouvelle séquence protéique correspondant à une chaîne de globine, notée A1.

Une protéine préférée selon l'invention est une protéine telle que définie ci-dessus, caractérisée en ce qu'elle comprend ou est constituée par :

– la séquence SEQ ID NO : 8,

15 – ou toute séquence dérivée de la séquence SEQ ID NO : 8 ou d'un fragment défini ci-dessous, notamment par substitution, suppression ou addition d'un ou plusieurs acides aminés, sous réserve que ladite séquence dérivée permette le transport de l'oxygène,

20 – ou toute séquence homologue de la séquence SEQ ID NO : 8 ou d'un fragment défini ci-dessous, ayant de préférence une homologie d'au moins environ 75%, notamment d'au moins environ 85%, avec la séquence SEQ ID NO : 8, sous réserve que ladite séquence homologue permette le transport de l'oxygène,

25 – ou tout fragment d'une des séquences définies ci-dessus, sous réserve que ledit fragment permette le transport de l'oxygène, notamment tout fragment étant constitué d'au moins environ 60 acides aminés, et notamment d'au moins environ 160 acides aminés contigus dans la séquence SEQ ID NO : 8.

La séquence SEQ ID NO : 8 est une nouvelle séquence protéique correspondant à une chaîne de globine, notée B2.

La présente invention concerne également des séquences nucléotidiques telles qu'obtenues selon le procédé tel que défini ci-dessus.

30 La présente invention concerne également des séquences nucléotidiques codant pour une protéine telle que définie ci-dessus.

La présente invention concerne également une séquence nucléotidique telle que définie ci-dessus, caractérisée en ce qu'elle comprend ou est constituée par :

– la séquence nucléotidique SEQ ID NO : 1 codant pour SEQ ID NO : 2,

— ou toute séquence nucléotidique dérivée, par dégénérescence du code génétique, de la séquence SEQ ID NO : 1, et codant pour une protéine représentée par SEQ ID NO : 2,

5 — ou toute séquence nucléotidique dérivée, notamment par substitution, suppression ou addition d'un ou plusieurs nucléotides, de la séquence SEQ ID NO : 1 codant pour une protéine dérivée de SEQ ID NO : 2,

— ou toute séquence nucléotidique homologue de SEQ ID NO : 1, ayant de préférence une homologie d'au moins environ 60%, avec la séquence SEQ ID NO : 1,

10 — ou tout fragment de la séquence nucléotidique SEQ ID NO : 1 ou des séquences nucléotidiques définies ci-dessus, ledit fragment étant de préférence constitué d'au moins environ 180 nucléotides, et notamment d'au moins environ 480 nucléotides contigus dans ladite séquence,

— ou toute séquence nucléotidique complémentaire des séquences ou fragments susmentionnés,

15 — ou toute séquence nucléotidique susceptible d'hybrider dans des conditions stringentes avec la séquence complémentaire d'une des séquences ou d'un des fragments susmentionnés.

Les conditions de stringence correspondent à des gammes de températures comprises entre 48 et 60°C et de concentrations en $MgCl_2$ comprises entre 1 et 3 mM.

20 La présente invention concerne également une séquence nucléotidique telle que définie ci-dessus, caractérisée en ce qu'elle comprend ou est constituée par :

— la séquence nucléotidique SEQ ID NO : 3 codant pour SEQ ID NO : 4,

25 — ou toute séquence nucléotidique dérivée, par dégénérescence du code génétique, de la séquence SEQ ID NO : 3, et codant pour une protéine représentée par SEQ ID NO : 4,

— ou toute séquence nucléotidique dérivée, notamment par substitution, suppression ou addition d'un ou plusieurs nucléotides, de la séquence SEQ ID NO : 3 codant pour une protéine dérivée de SEQ ID NO : 4,

30 — ou toute séquence nucléotidique homologue de SEQ ID NO : 3, ayant de préférence une homologie d'au moins environ 60%, avec la séquence SEQ ID NO : 3,

— ou tout fragment de la séquence nucléotidique SEQ ID NO : 3 ou des séquences nucléotidiques définies ci-dessus, ledit fragment étant de préférence constitué

d'au moins environ 180 nucléotides, et notamment d'au moins environ 480 nucléotides contigus dans ladite séquence,

- ou toute séquence nucléotidique complémentaire des séquences ou fragments susmentionnés,

- ou toute séquence nucléotidique susceptible d'hybrider dans des conditions stringentes avec la séquence complémentaire d'une des séquences ou d'un des fragments susmentionnés.

La présente invention concerne également une séquence nucléotidique telle que définie ci-dessus, caractérisée en ce qu'elle comprend ou est constituée par :

- la séquence nucléotidique SEQ ID NO : 5 codant pour SEQ ID NO : 6,
- ou toute séquence nucléotidique dérivée, par dégénérescence du code génétique, de la séquence SEQ ID NO : 5, et codant pour une protéine représentée par SEQ ID NO : 6,

- ou toute séquence nucléotidique dérivée, notamment par substitution, suppression ou addition d'un ou plusieurs nucléotides, de la séquence SEQ ID NO : 5 codant pour une protéine dérivée de SEQ ID NO : 6,

- ou toute séquence nucléotidique homologue de SEQ ID NO : 5, ayant de préférence une homologie d'au moins environ 60%, avec la séquence SEQ ID NO : 5,

- ou tout fragment de la séquence nucléotidique SEQ ID NO : 5 ou des séquences nucléotidiques définies ci-dessus, ledit fragment étant de préférence constitué d'au moins environ 180 nucléotides, et notamment d'au moins environ 480 nucléotides contigus dans ladite séquence,

- ou toute séquence nucléotidique complémentaire des séquences ou fragments susmentionnés,

- ou toute séquence nucléotidique susceptible d'hybrider dans des conditions stringentes avec la séquence complémentaire d'une des séquences ou d'un des fragments susmentionnés.

La présente invention concerne également une séquence nucléotidique telle que définie ci-dessus, caractérisée en ce qu'elle comprend ou est constituée par :

- la séquence nucléotidique SEQ ID NO : 7 codant pour SEQ ID NO : 8,

- ou toute séquence nucléotidique dérivée, par dégénérescence du code génétique, de la séquence SEQ ID NO : 7, et codant pour une protéine représentée par SEQ ID NO : 8,

— ou toute séquence nucléotidique dérivée, notamment par substitution, suppression ou addition d'un ou plusieurs nucléotides, de la séquence SEQ ID NO : 7 codant pour une protéine dérivée de SEQ ID NO : 8,

5 — ou toute séquence nucléotidique homologue de SEQ ID NO : 7, ayant de préférence une homologie d'au moins environ 60%, avec la séquence SEQ ID NO : 7,

— ou tout fragment de la séquence nucléotidique SEQ ID NO : 7 ou des séquences nucléotidiques définies ci-dessus, ledit fragment étant de préférence constitué d'au moins environ 180 nucléotides, et notamment d'au moins environ 480 nucléotides contigus dans ladite séquence,

10 — ou toute séquence nucléotidique complémentaire des séquences ou fragments susmentionnés,

— ou toute séquence nucléotidique susceptible d'hybrider dans des conditions stringentes avec la séquence complémentaire d'une des séquences ou d'un des fragments susmentionnés.

15 La présente invention concerne un procédé de préparation tel que défini ci-dessus, de séquences nucléotidiques codant pour les chaînes protéiques constituant la molécule d'hémoglobine d'Annélides, notamment d'*Arenicola marina*, ledit procédé étant caractérisé en ce qu'il comprend les étapes suivantes :

20 — une étape de mise en présence de la molécule d'hémoglobine susmentionnée avec au moins un agent dissociant, notamment un mélange constitué de dithiothréitol (DTT) ou de chlorhydrate de tris(2-carboxyéthyl)phosphine (TCEP) ou de bêta-mercaptoéthanol et d'un tampon de dissociation, pendant un temps suffisant pour séparer les chaînes protéiques les unes des autres,

25 permettant la dissociation de ladite molécule d'hémoglobine, afin d'obtenir les dix chaînes protéiques constituant ladite molécule,

— l'isolement de chacune des dix chaînes protéiques susmentionnées,

— le microséquençage de chacune des dix chaînes protéiques isolées susmentionnées par spectrométrie de masse, afin d'obtenir une microséquence correspondant à chacune des dix séquences constituées de 5 à 20 acides aminés,

30 — la détermination des dix couples d'amorces dégénérées (sens et antisens) à partir des microséquences susmentionnées,

— la préparation des séquences nucléotidiques codant pour les chaînes protéiques susmentionnées, à partir des amorces telles qu'obtenues précédemment, par un procédé d'amplification en chaîne par polymérase (PCR), comprenant les étapes suivantes :

- la première étape dudit procédé est une étape de dénaturation de l'ADNc codant pour chacune des dix chaînes protéiques constituant la molécule d'hémoglobine d'*Arenicola marina*, d'environ 10 secondes à environ 5 minutes à une température comprise d'environ 90°C à environ 110°C,

- le cycle, répété environ 30 à 40 fois, comprend les étapes suivantes :

* une étape de dénaturation de l'ADNc codant pour chacune des dix chaînes protéiques constituant la molécule d'hémoglobine d'*Arenicola marina*, d'environ 10 secondes à environ 5 minutes, à une température comprise d'environ 90°C à environ 110°C,

* une étape d'hybridation des couples d'amorces de l'invention aux brins d'ADNc monocaténares susmentionnés pour obtenir des amorces hybridées, d'environ 20 secondes à environ 2 minutes, à une température comprise d'environ 50°C à environ 56°C,

* une étape d'élongation des amorces hybridées telles qu'obtenues précédemment par une polymérase d'environ 20 secondes à environ 1 minute et 30 secondes, à une température comprise d'environ 70°C à environ 75°C, et

- la dernière étape du procédé est une étape d'élongation des amorces hybridées telles qu'obtenues précédemment par une polymérase d'environ 5 minutes à environ 15 minutes à une température comprise d'environ 70°C à environ 75°C.

DESCRIPTION DES FIGURES

La Figure 1 représente un chromatogramme de l'hémoglobine de *Arenicola marina* sur colonne Superose 12-C. La courbe supérieure correspond à une absorbance de 414 nm et la courbe inférieure à une absorbance de 280 nm. (Le collecteur est programmé pour collecter entre 16 et 18 minutes).

La Figure 2 représente le spectre UV de l'hémoglobine de *Arenicola marina* fonctionnelle (sous sa forme oxyhémoglobine).

La Figure 3 représente le chromatogramme (à 414 nm) de l'HbAm dissociée (partiellement) obtenu sur Superose 12-C et les traits verticaux sur le chromatogramme correspondent aux fenêtres de collecte (correspondant à la récupération des sous-unités).

La Figure 4 représente un gel SDS-PAGE obtenu pour les différentes fractions collectées.

La Figure 5 représente le chromatogramme (à 414 nm) de l'HbAm dissociée (partiellement) obtenu sur CIM DISK DEAE (système d'échange anionique) et les traits verticaux sur le chromatogramme correspondent aux fenêtres de collecte. La courbe en pointillés indique le gradient.

La Figure 6 représente un gel SDS-PAGE obtenu pour les différentes fractions collectées.

La Figure 7 représente la cinétique de dissociation de l'HbAm en présence d'urée 3M. L'axe des abscisses correspond au nombre de jours et l'axe des ordonnées correspond au pourcentage de dissociation de la molécule native ; la courbe en pointillés correspond au dodécamère ; la courbe avec les carrés noirs au trimère et au "linker" (chaîne de structure) ; la courbe avec les ronds noirs aux monomères.

La Figure 8 représente la cinétique de dissociation de l'HbAm à pH 10. L'axe des abscisses correspond au nombre de jours et l'axe des ordonnées correspond au

pourcentage de dissociation de la molécule native ; la courbe en pointillés correspond au dodécamère ; la courbe avec les carrés noirs au trimère et au "linker" ; la courbe avec les ronds noirs aux monomères.

5 La Figure 9 représente la cinétique de dissociation de l'HbAm en présence d'urée 3M à pH 10. L'axe des abscisses correspond au nombre de jours et l'axe des ordonnées correspond au pourcentage de dissociation de la molécule native ; la courbe en pointillés correspond au dodécamère ; la courbe avec les carrés noirs au trimère et au "linker" ; la courbe avec les ronds noirs aux monomères.

10 La Figure 10 représente le suivi de la cinétique de réassociation à partir du pourcentage d'HbAm (HBL) et de Dodécamère (D) et selon la technique de changement de tampon (Centricon ou Dialyse). L'axe des abscisses correspond au nombre de jours et l'axe des ordonnées correspond au pourcentage de dissociation de la molécule native avec la technique Centricon ; la courbe en traits pointillés correspond à HBL avec la technique de dialyse ; la courbe avec les triangles noirs correspond au dodécamère avec la technique Centricon ; la courbe en trait plein correspond au dodécamère avec la technique de dialyse.

20 La Figure 11 représente la superposition des chromatogrammes de chromatographie d'exclusion au cours de la réassociation correspondant à différents temps de réassociation.

25 La Figure 12 représente le chromatogramme HPLC obtenu avec les fenêtres de collecte (traits verticaux). Les masses indiquées ont été déterminées par spectrométrie de masse (ESI-TOF). Les codes (d2, a1, b3, d1, c) correspondent aux noms des globines telles que mentionnées dans l'article de Zal et al. (1997).

PARTIE EXPÉRIMENTALE

La présente invention a pour objectif l'utilisation de l'hémoglobine extracellulaire du polychète marin *Arenicola marina* (HbAm) comme substitut sanguin chez des vertébrés. Cependant, même si ce ver représente une biomasse importante, la synthèse par génie génétique s'avère une voie indispensable et nécessaire. Il est donc primordial d'étudier les propriétés d'auto-assemblage de cette molécule afin de développer une hémoglobine artificielle, fonctionnelle et stable. Les protocoles de dissociations ainsi que les techniques de purification de chaque sous-unité et chaîne polypeptidique sont développées en détail ci-après, de même que les premières approches de réassociation.

Extraction et purification de l'hémoglobine

1) Espèce étudiée : *Arenicola marina* ; Annélide de l'écosystème intertidal

L'Annélide Polychète *Arenicola marina* est une espèce sédentaire très répandue sur toutes les côtes de l'Atlantique Nord, de la mer Noire et de l'Adriatique situées au-dessus du quarantième parallèle. Dans la région de Roscoff, l'Arénicole, communément appelée "ver du pêcheur" constitue des populations denses. Le sédiment habité par ces populations présente une surface irrégulière alternant bosses et creux formés respectivement par des monticules de déjections et des dépressions coniques. L'Arénicole vit dans des galeries aménagées dans le sable. La structure de la galerie se présente sous forme de U, avec une branche ouverte sur l'extérieur, l'autre étant fermée. L'Arénicole est logée dans la partie horizontale, son extrémité céphalique orientée vers la partie aveugle. Elle ingère le sable, en extrait la matière organique assimilable et défèque ensuite par son extrémité caudale, formant ainsi des monticules de tortillons de sable. A l'intérieur de l'étage médiolittoral, la répartition et la densité des peuplements sont essentiellement commandées par la granulométrie, la concentration en matière organique et la salinité.

L'Arénicole, vivant surtout dans les zones de balancement des marées, est amenée à subir des variations de pression d'oxygène. Sa galerie lui permet d'être en contact permanent avec l'eau de mer (riche en oxygène), pendant la marée basse.

2) Méthodes d'étude

2.1. Prélèvement du matériel biologique

Les animaux sont récoltés à marée basse, dans la baie de Penpoull, près de Roscoff (France) et maintenus dans l'eau de mer une nuit afin de vider leur tube digestif. Les prélèvements de sang sont effectués à l'aide d'une seringue au niveau du vaisseau dorsal. Les échantillons sont collectés sur de la glace et filtrés sur de la laine de verre. Après une centrifugation à froid (15 000 g pendant 15 min à 4°C) pour éviter la dissociation des molécules et pour éliminer les débris tissulaires, les surnageants sont concentrés au moyen d'une cellule Amicon (Millipore) et d'une membrane de "cut-off" de 500 kDa (ne sont retenues que les masses supérieures à 500 kDa).

2.2. Purification des hémoglobines

Une fois le sang concentré, une filtration basse pression (FPLC) par exclusion (séparation en fonction de la taille de la molécule : plus la molécule est de taille importante plus elle est éluée rapidement) est effectuée sur colonne (100 × 3 cm) de gel Sephacryl S-400 (Amersham) (gamme de séparation comprise entre 20×10^3 et 8000×10^3), en chambre froide (4°C). Chaque purification est effectuée sur 5 mL d'échantillon, élués avec le tampon salé *Arenicola marina* (10 mM Hepes ; 4 mM KCl ; 145 mM NaCl ; 0,2 mM MgCl₂ ajusté à pH 7,5 avec de l'acide chlorhydrique). Le débit utilisé pour cette première purification est de 40 tr/min et ne sont récupérées que les fractions les plus rouges (contenant de l'hème). Ces fractions sont ensuite concentrées à l'aide de tube Centricon-10kDa retenant les molécules d'un poids supérieur ou égal à 10000 Da.

Une deuxième purification est ensuite effectuée par filtration basse pression (Système HPLC, Waters) d'aliquot de 200 µL sur une colonne 1 × 30 cm Superose 12-C (Pharmacia, gamme de séparation comprise entre 5×10^3 et 3×10^5 Da) à température ambiante. Le débit utilisé est de 0,5 ml/min. Les échantillons sont conservés à 4°C et récoltés dans de la glace. L'absorbance de l'éluat est suivie à deux longueurs d'ondes : 280 nm (pic d'absorbance des protéines) et 414 nm (pic d'absorbance de l'hémoglobine). Les fractions contenant de l'hème sont isolées grâce au collecteur (programmé sur une fenêtre de temps correspondant au temps de rétention de l'hémoglobine) (Figure 1). Les échantillons sont concentrés, dosés, puis stockés à -40°C avant utilisation.

2.3. Dosage des hémoglobines

Le réactif de Drabkin (Sigma), utilisé pour le dosage, permet de déterminer la quantité en hème de la solution. L'hémoglobine réagit avec le réactif de Drabkin lequel contient du potassium de ferricyanure, du cyanure de potassium et du bicarbonate de sodium. L'hémoglobine est transformée en méthémoglobine par l'action du ferricyanure. Les méthémoglobines réagissent ensuite avec le cyanure pour former le cyanmethémoglobine. L'absorbance de ce dérivé à 540 nm est proportionnelle à la quantité d'hème dans la solution. Or, l'hémoglobine extracellulaire de *Arenicola marina* (HBL) contient en moyenne 1 mol d'hème pour 23 000 g de protéine, ce qui permet par un calcul simple, d'obtenir la concentration en HBL de chaque échantillon.

3) Résultats

Ainsi, 800 mg d'hémoglobine extracellulaire de *Arenicola marina* ont été purifiés. Chaque lot (aliquots de 1mL) est analysé par FPLC sur colonne Superose 12-C (Pharmacia) afin de s'assurer de la pureté de l'échantillon (un pic unique). De même un spectre UV sur une plage de 400 nm à 700 nm est réalisé pour vérifier la fonctionnalité des hémoglobines de chaque lot (Figure 2). On observe deux maximums d'absorption à 541 et 577 nm. Par comparaison, on rappelle que la méthémoglobine présente un maximum à 540 nm.

Enfin, ces lots sont utilisés par la société Biotrial S.A. (Rennes, France) pour être transfusés à des souris et tester les éventuelles réactions pathologiques et immunogènes.

Dissociation de l'hémoglobine extracellulaire de *Arenicola marina* en ces différentes sous-unités de base (Trimères, Dimères de linker et Monomères)

La dissociation de l'hémoglobine extracellulaire d'*Arenicola marina* (HbAm) doit être totale et conserver les sous-unités fonctionnelles. Les différentes sous-unités sont ensuite isolées et analysées par des techniques de chromatographie liquide (exclusion et échange d'ions), développées à cet effet.

1) Dissociation de l'HbL

1.1. Technique de dissociation

Les études préliminaires de dissociation ont été développées à partir des publications de l'art antérieur (Vinogradov et al., 1979 ; Sharma et al., 1996 ; Mainwaring et al., 1986 ; Polidori et al., 1984 ; Kapp et al., 1984 ; Chiancone et al., 1972 ; Vinogradov et al., 1991 ; Krebs et al., 1996), c'est-à-dire en présence d'un réactif dissociant unique : urée, ions hétéropolytungstate ou ions hydroxydes. Les trois agents utilisés agissent différemment sur la molécule :

- les ions hydroxydes (OH^-) déstabilisent les ponts salins
- l'urée déstabilise les interactions hydrophobes
- l'action des ions polytungstate n'est pas encore bien connue mais ils interagissent sur les interactions électrostatiques.

Or, l'intérêt est d'obtenir le plus rapidement et le plus efficacement possible les quatre sous-unités de bases d'où l'idée d'associer les différents agents dissociants et en particulier le pH alcalin et l'urée. Après différents tests, il s'avère qu'une dissociation rapide et efficace est obtenue avec de l'urée 3M dilué dans le tampon de dissociation (0,1 M de base Trisma et 1mM d'EDTA) ajusté à pH 10 avec de la soude 2N. L'HbAm est ajustée à une concentration d'environ 4 mg/mL (solution mère). Toutes les analyses sont effectuées à $+4^\circ\text{C}$ et les échantillons sont conservés à l'obscurité le temps de l'étude. (Trisma = tris[hydroxyméthyl]aminométhane)

1.2. Analyses par chromatographie d'exclusion

Les conditions d'analyses sont détaillées dans le tableau 1 ci-dessous.

Système HPLC	HPLC Waters 626 LC System	
Colonne	Superose 12-C (Pharmacia) (gamme de séparation comprise entre 5×10^3 et 3×10^5 Da)	
Débit	0.5 mL/min	
Eluant	tampon de dissociation cité plus haut ajusté à pH 7,0 avec de l'acide chlorhydrique concentré et filtré sur filtre 0,45 μm	
Température de l'injecteur	$+4^\circ\text{C}$	
	Echantillon	
	Suivi Analytique	Séparation et collecte
Volume injecté	20 μL	200 μL
Préparation des échantillons	Solution mère diluée à 1mg/mL dans le tampon de dissociation à pH10 et filtrée sur 0,45 μm	Solution mère filtrée sur filtre 0,45 μm

1.3. Analyses par échange d'ions

Le point isoélectrique (pHi) de l'HBL étant de 4,69 (Vinogradov, 1985), l'analyse par échange d'ions est réalisée sur une colonne anionique CIM DEAE disk (Interchim). En effet, L'HBL est chargée négativement pour un pH supérieur au pHi et est donc fixée sur une résine chargée positivement (résine DEAE). L'élution est effectuée grâce à la force ionique avec un gradient non linéaire de NaCl de 0 à 1 M (solution de NaCl de 1 M diluée dans le tampon de dissociation à pH 7,0 et filtrée sur 0,45 µm). Le tampon de dissociation à pH 7 est utilisé comme tampon d'élution. Le débit est de 4 mL/min.

2) Résultats

2.1. Chromatographie d'exclusion

Le chromatogramme de l'HbAm dissociée en partie, est représenté figure 3. Cinq pics sont observés et il s'agit de les identifier. Les molécules sont éluées selon leur masse décroissante et l'HbAm native sort dans le volume mort (16 min). Les fractions correspondant à chaque pic sont analysées par SDS-PAGE (figure 4). Les résultats sont présentés dans le tableau 2 ci-dessous.

Fractions	Temps de rétention	Sous-unités
1	16 min 40	HbAm native
2	22 min 20	Dodécamère
3	25 min 30	Dimère de linker
4	26 min 40	Trimère
5	28 min 30	Monomères

2.2. Chromatographie ionique

Une fois la méthode développée (tableau 3), les fractions sont collectées, concentrées et analysées par gel SDS afin d'identifier chaque pic (Fig. 6 et tableau 4).

Une méthode est ensuite développée pour repurifier chaque sous-unité.

Temps en min	A : 1M NaCl dans B	B : Tampon de dissociation pH 7.0
0	5%	95 %
0,5	15%	85 %
1,5	15%	85 %
2,5	22%	78 %
3,5	22%	78 %
3,6	25%	75 %
5,5	25%	75 %
5,6	29%	71 %
6,5	29%	71 %
6,6	36%	64 %
7,0	36%	64 %
8,0	45%	55 %
8,1	100%	0 %
9,5	100%	0 %

TABLEAU 3 : Méthode développée pour l'analyse de l'HbL dissociée sur CIM-DEAE

fractions	Temps de rétention	Sous-unités
1	1 min 15	Monomères
2	2 min 40	Dimère de linker
3	3 min 10	Dimère de linker
4	4 min 40	Dodécamère
5	6 min 30	?
6	7 min 30	Trimère
7	8 min 50	HbAm

TABLEAU 4 : Association réalisée après analyse du gel (Figure 6) entre le temps de rétention et les sous-unités.

2.3. Dissociation de l'HbAm

La cinétique de dissociation est suivie par chromatographie d'exclusion. L'intégration des chromatogrammes par le logiciel Millénium (Waters), permet le calcul du pourcentage des différents composés à partir de l'aire sous courbe. L'évolution de la cinétique de dissociation est représentée figures 7, 8 et 9.

Les trois graphiques des Figures 7, 8 et 9 montrent bien l'intérêt d'associer les deux agents dissociants (urée 3M et OH^-) pour obtenir efficacement les trois sous-unités de base en 24h.

5

Réassociation de l'hémoglobine

1) Matériels et méthodes

Les expériences de réassociation sont effectuées sur l'HbAm dissociée selon les protocoles cités plus haut (pH9, pH10, Urée 3M, Urée 4M, Urée 3M à pH10). Différents tampons de réassociation sont testés afin d'obtenir une réassociation optimale. Le changement de tampon (tampon de dissociation \rightarrow tampon de réassociation) est effectué de deux façon différentes pour tester la "meilleure" technique de remplacement de tampon :

- 15 – L'HbAm dissociée est lavée 4 fois contre 4 mL de tampon de réassociation sur Centricon-10 (Millipore) à $+4^\circ\text{C}$;
- L'HbAm dissociée est dialysée 24 h contre de l'eau MilliQ (Millipore) ($2 \times 2\text{L}$) puis 48 h contre le tampon de réassociation ($3 \times 2\text{L}$) à $+4^\circ\text{C}$.

20

2) Résultats

D'après des résultats ultérieurs sur les hémoglobines extracellulaires d'Annélides (Mainwaring et al., 1986 ; Polidori et al., 1984), la présence d'ions divalents tels que Ca^{2+} et Mg^{2+} est nécessaire au maintien de la structure quaternaire de l'hémoglobine. En effet, ils la stabilisent et ralentissent le phénomène de dissociation (Sharma et al., 1996).
25 Ces ions forment un complexe avec les groupes carboxylates des chaînes latérales et carbonyles des chaînes principales. La présence d'ions divalents peut avoir un effet sur la réassociation lorsque les groupements carboxyliques sont ionisés, donc entre autre lorsque la dissociation a eu lieu à pH alcalin. Il est donc important que le tampon de réassociation contienne du calcium et/ou du magnésium. Ceci explique aussi la présence
30 d'EDTA dans le tampon de dissociation ; EDTA qui chélate ces ions divalents. Le tampon actuellement mis au point est constitué de 0,1 M de base Trisma, 10 mM de CaCl_2 et 10 mM de MgCl_2 ajusté à pH 7 avec de l'HCl concentré. La réassociation est suivie selon le même principe que la dissociation (Figures 10 et 11).

Un début de réassociation est observé après 15 jours et 40% de l'HbAm sont réassociés après 45 jours dans le tampon de réassociation.

Réduction de l'HbAm pour l'étude des différentes chaînes polypeptidiques

1) Réduction de l'hémoglobine

L'HbAm (4 mg/mL) est réduite dans 100 mM de DTT (dithiothréitol) dissout dans le tampon de dissociation à pH 8-9 pendant 1 h à 40°C. Dans ces conditions drastiques, seules les globines pourront être analysées. En effet, les linkers (protéines non-globines) sont riches en cystéines et sont donc détériorés. Une fois réduite, l'HbAm est lavée 4 fois sur Amicon 30 000Da (Millipore) dont seul le filtrat est récupéré (tout ce qui est de poids inférieur à 30 000 Da). Le filtrat est ensuite lavé sur Amicon 10 000Da pour éliminer tout ce qui est inférieur à 10 000 Da. Ainsi, seuls les monomères compris entre 30 000 Da et 10 000 Da sont contenus dans l'échantillon (gamme de poids des chaînes de globines qui constituent l'HbAm).

2) Purification des globines par chromatographie de phase inverse

2.1. Matériels et méthodes

La chromatographie par phase inverse est effectuée sur une colonne de Symmetry 300 C₁₈ 5µm (4,6 × 250 mm) de chez Waters. La méthode développée est décrite dans le tableau ci-dessous et le chromatogramme obtenu est représenté figure 12.

débit	1 mL/min	
Eluant A	H ₂ O MilliQ + 0.1% v/v TFA	
Eluant B	ACN + 0.1% v/v TFA	
Gradient		
Temps en min	% A	% B
0	80	20
0,25	60	40
4,0	55	45
10,0	55	45
10,05	0	100
16,0	0	100

3) Analyse par micro séquençage des globines isolées

Les chaînes de globines isolées sont microséquencées par spectrométrie de masse afin d'obtenir des informations de séquences pour réaliser des amorces ou pour compléter les séquences préalablement déterminées en biologie moléculaire.

5

Détails des protocoles d'amplification par PCR et présentation des séquences des globines des sous-familles A1, A2a, A2b et B2 du polychète marin *Arenicola marina*

10

Les amplifications par PCR des 4 globines A1, A2a, A2b et B2, dont les séquences nucléotidiques sont présentées ci-dessous, ont débuté par la conception d'amorces dégénérées spécifiques (sens et antisens) des sous-familles A1, A2 et B2. Ces amorces ont été conçues à partir d'alignements de séquences protéiques de globines d'annélides disponibles dans les banques de données.

15

Préparation des microséquences par MSMS (Séparation des protéines sur gel 2D)

Electrofocalisation

20

Après purification par FPLC, l'hémoglobine de *Arenicola* est dialysée et lyophilisée. 500 µg sont repris dans un tampon de réhydratation. Ce tampon contient du Chaps (3-[(3-cholamidopropyl)diméthylammonio]-1-propane-sulfonate ; Sigma) à 4%, 1% de DTT fraîchement préparé, cocktail d'inhibiteurs de protéase (Bohringer), 50 µg/ml de TLCK (inhibiteur de la trypsine), 1 µl de Bleu de Bromophénol à 1%.

25

Le mélange est soniqué et centrifugé pour éliminer le matériel non dissous. On dépose ensuite de l'huile de roche aux deux extrémités du support de la bande d'électrofocalisation, et on dépose ensuite l'échantillon au milieu. La bande de 17 cm est ensuite déposée sur l'échantillon en éliminant toute bulle d'air. La bande est alors recouverte d'huile de roche pour éviter l'évaporation de l'échantillon.

30

Une réhydratation active est alors réalisée à 50V (20°C durant 12 heures). Ensuite nous procédons à la focalisation après avoir recouvert la bande de 2 papiers wiks (Biorad) mouillés en contact avec les deux électrodes durant deux jours.

La bande sera ensuite récupérée et placée sur un gel en gradient d'acrylamide 6-18%, de largeur 18 cm, de longueur 20 cm et d'épaisseur 1 mm. Ce gel a été coulé à un

débit constant de 8,4 ml/min, à l'aide d'un formeur de gradient. La séparation des protéines sera alors réalisée en fonction de leur taille après avoir scellé la bande en haut du gel à l'aide d'une solution d'agarose à 1%. Une fois séparées, les bandes protéiques sont révélées sur gel par coloration au bleu de Coomassie (Coomassie® G250).

Construction du gel en gradient

Vingt-cinq ml de solution dense à 18% de polyacrylamide (acrylamide 2,5 M ; Tris 0,4 M ; Glycérol 30% (v/v) ; dodécyl sulfate de sodium (SDS) 3,5 mM ; TEMED (N,N,N',N'-tétraméthyléthylènediamine ; Sigma) 0,05% (v/v) ; persulfate de sodium 1,6 mM) sont placés dans la chambre de mélange sous agitation constante tandis qu'un même volume de solution légère à 6% de polyacrylamide (acrylamide 0,8 M ; Tris 0,4 M ; SDS 3,5 mM ; TEMED 0,06% (v/v) ; persulfate de sodium 2,4 mM) est placé dans l'autre chambre. Le haut du gel est recouvert d'une solution d'isobutanol saturée en eau bidistillée. Le gel est ensuite laissé 1 h à polymériser à température ambiante puis le haut du gel est rincé plusieurs fois avec de l'eau bidistillée et l'ensemble est placé une nuit à 10°C. Après avoir enlevé l'eau résiduelle à l'aide d'un papier absorbant, la solution de gel de concentration (acrylamide 0,56 M ; méthylène bis-acrylamide 6,9 mM ; Tris 124 mM ; SDS 3,5 mM ; TEMED 0,05% (v/v) ; persulfate de sodium 2,2 mM) est coulée sur le gel de séparation et une cale permettant de former l'empreinte de la bande est introduite dans la solution gel de concentration. La polymérisation est complète après 1 h à température ambiante.

Migration électrophorétique des protéines.

La migration est effectuée en enceinte réfrigérée à 10°C, pendant 14h à 400 V, 25 mA et 100 W.

Coloration au bleu de Coomassie colloïdal

Le gel de polyacrylamide est tout d'abord lavé à l'eau distillée 3 × 30 minutes.

Le gel est ensuite placé pendant une heure dans un récipient en inox contenant 150 ml de solution commerciale de bleu de Coomassie, sous agitation douce, à température ambiante. Après quatre lavages à l'eau bidistillée, le gel est finalement conservé à 4°C dans l'eau stérile jusqu'à l'excision des bandes polypeptidiques.

Digestion des protéines séparées sur gel 1D

Chaque bande polypeptidique du gel est découpée de manière systématique tous les 2 mm sur toute leur longueur. Chaque morceau de gel ainsi obtenu est soumis à une digestion enzymatique. Cette étape de digestion enzymatique est essentielle. Elle consiste à hydrolyser les protéines de façon spécifique, à l'aide d'une enzyme, en plusieurs peptides, permettant ainsi leurs extractions du morceau de gel.

Etape de décoloration, réduction, alkylation

Avant de débiter la digestion, des étapes de décoloration, de réduction et d'alkylation sont indispensables :

- les lavages successifs à l'hydrogencarbonate d'ammonium (NH_4HCO_3) et à l'acétonitrile (ACN) permettent d'éliminer l'agent de coloration présent dans le morceau de gel,
- les réactions de réduction au dithiothréitol (DTT) et d'alkylation à l'iodoacétamide permettent l'ouverture puis le blocage des ponts disulfures formés entre deux cystéines présentes dans la séquence de la protéine et des liaisons cystéine-acrylamide.

Etape de digestion

L'étape de digestion proprement dite, utilise une endoprotéase, la trypsine (25 kDa). Cette enzyme coupe la protéine par une hydrolyse au niveau C-terminal de la lysine et de l'arginine, donnant généralement des peptides de masses comprises entre 500 et 2500 Da.

Etape d'extraction

La dernière étape de la méthode consiste à extraire du gel les peptides tryptiques à l'aide d'une solution d'extraction, composée d'acétonitrile et d'eau, additionnée d'un peu d'acide.

Analyse par nanoLC-MS-MS

Les peptides extraits sont ensuite transférés vers la plaque PCR. Ce transfert se fait en deux fois 15 μL , pour récupérer la totalité du volume. Pour éliminer l'acétonitrile, qui pourrait gêner la rétention des peptides sur la pré-colonne, un temps d'évaporation (pause) de 2 heures est appliqué avant l'analyse par nanoLC-MS-MS.

Les matrices d'ADN complémentaires utilisées pour les réactions de PCR ont été synthétisées à partir d'ARN messagers purifiés à partir d'ARN totaux extraits sur des Arénicoles, en raison de la petite taille des organismes et de leur rythme intense de croissance traduisant des niveaux importants d'expression des gènes, dont ceux impliqués dans la synthèse de l'hémoglobine. Les ADN complémentaires ont ainsi été synthétisés. Ces étapes ont fait appel à des kits commerciaux de biologie moléculaire de chez Ambion (purification des ARNs), Amersham (purification des ARNm), Promega (RT), Invitrogene (clonage), Abgene (séquençage).

Dans un second temps, nous avons mis au point les réactions de PCR, notamment en ce qui concerne la détermination des paramètres de temps de dénaturation, de temps et de température d'hybridation et de temps d'élongation. Les concentrations en $MgCl_2$ ont également été optimisées.

Sont présentées ci-dessous, les séquences nucléotidiques des amorces dégénérées sens et antisens, les paramètres des PCR et les séquences codantes partielles pour chacune des globines A2a, A2b, A1 et B2.

Pour les chaînes de structures, on utilise les amorces susmentionnées SEQ ID NO : 17 et SEQ ID NO : 18.

On précise que l'analyse totale de ces séquences par un blast dans les banques de données donne des valeurs comprises entre $2.10^{-3} < \text{Evalue} < 5^{-31}$.

Globine A2 (a)

Pour obtenir la séquence nucléotidique SEQ ID NO : 1 codant pour la globine A2(a) (SEQ ID NO : 2), on utilise le couple d'amorces (SEQ ID NO : 9 ; SEQ ID NO : 10).

Les conditions de PCR sont les suivantes :

Temps et température initiale de dénaturation :	4 min à 95°C	} 35 cycles
Temps et Température de dénaturation :	30 s à 95°C	
Temps et Température d'hybridation :	30 s à 56°C	
Temps et Température d'élongation :	40 s à 72°C	

Temps et Température d'élongation finale : 10 min à 72°C

Réaction de PCR :

5 Par réaction : 1 µL ADNc
100 ng amorce sens
100 ng amorce antisens
dNTP 200 µM final
MgCl₂ 2mM final
10 Tampon PCR 1X final
1 unité Taq Polymérase
Qsp 25 µL H₂O

15 Globine A2 (b)

Pour obtenir la séquence nucléotidique SEQ ID NO : 3 codant pour la globine A2(b) (SEQ ID NO : 4), on utilise le couple d'amorces (SEQ ID NO : 11 ; SEQ ID NO : 12).

20

Les conditions de PCR sont les suivantes :

PCR (1) : 4 min à 95°C

30 s à 95°C

30 s à 52°C

40 s à 72°C

10 min à 72°C

35 cycles

PCR (2) : 4 min à 95°C

30 s à 95°C

30 s à 52°C

40 s à 72°C

10 min à 72°C

35 cycles

Globine A1

30

Pour obtenir la séquence nucléotidique SEQ ID NO : 5 codant pour la globine A1 (SEQ ID NO : 6), on utilise le couple d'amorces (SEQ ID NO : 13 ; SEQ ID NO : 14).

Les conditions de PCR sont les suivantes :

5 PCR : 4 min à 95°C
 1 min à 95°C
 1 min à 50°C } 35 cycles
 1 min 30 à 72°C
 10 min à 72°C

10 Globine B2

Pour obtenir la séquence nucléotidique SEQ ID NO : 7 codant pour la globine B2 (SEQ ID NO : 8), on utilise le couple d'amorces (SEQ ID NO : 15 ; SEQ ID NO : 16):

15 Les conditions de PCR sont les suivantes :

20 PCR (1) : 4 min à 95°C
 30 s à 95°C
 40 s à 52°C } 35 cycles
 30 s à 72°C
 10 min à 72°C

PCR (2) : 4 min à 95°C
 30 s à 95°C
 40 s à 52°C } 35 cycles
 30 s à 72°C
 10 min à 72°C

REFERENCES BIBLIOGRAPHIQUES

- 5
- Bunn, H.F. et Jandl, J.H. (1968) *Trans Assoc Am Physicians*, **81**, 147,
 - Chang, T.M.S. (1957) Hemoglobin Corpuscles, McGill University,
 - Chang, T.M.S. (1964) *Science*, **146**, 524-525,
 - Chang, T.M.S. (1971) *Biochem. Biophys. Res. Com.*, **44**, 1531-1533,
 - Chang, T.M.S. (1997) Blood substitutes: principales, methods, products and clinical trials, Vol. I, Karger, Basel, Suisse,
- 10
- Chiancone, E., et al. (1972) Studies on erythrocrucorin. II. Dissociation of earthworm erythrocrucorin, *J Mol Biol*, **70**(1): 73-84,
 - Clark, L.C.J. et Gollan, F. (1966) *Science*, **152**, 1755,
 - De Haas, F.; Boisset, N.; Taveau, J.C.; Lambert, O.; Vinogradov, S.N. et Lamy, J.N. (1996) *Biophys. J.*, **70**, 1973-1984,
- 15
- De Haas, F.; Taveau, J.-C.; Boisset, N.; Lambert, O.; Vinogradov, S.N. et Lamy, J.N. (1996) *J. Mol. Biol.*, **255**, 140-153,
 - De Haas, F.; Zal, F.; Lallier, F.H.; Toulmond, A. et Lamy, J.N. (1996) Proteins-structure function and genetics, **3**, 241-256,
 - De Haas, F.; Zal, F.; You, V.; Lallier, F.H.; Toulmond, A. et Lamy, J.N. (1996) *J. Mol. Biol.*, **264**, 111-120,
- 20
- Dickerson, R.E. et Geis, I. (1983) Hemoglobin: Structure, function, evolution, and pathology., Benjamin/Cummings, Menlo Park, CA,
 - Geyer, R.P.; Monroe, R.G. et Taylor, K. (1968) Survival of rats totally perfused with a fluorocarbon-detergent preparation., J.C. Norman, J. Folkman, W.G. Hardisson, L.E. Rudolf et F.J. Veith (Eds), 85-96, Appleton Century Crofts, New York,
- 25
- Goodin, T.H.; Grossbard, E.B.; Kaufman, R.J.; Richard, T.J.; Kolata, R.J.; Allen, J.S. et Layton, T.E. (1994), *Crit Care Med*, **22**, 680-689,
 - Hirsch, R.E.; Jelicks, L.A.; Wittenberg, B.A.; Kaul, D.K.; Shear, H.L. et Harrington, J.P. (1997) Artificial Cells, Blood Substitutes & Immobilization Biotechnology, **25**, 429-444,
- 30
- Jia, L.; Bonaventura, C.; Bonaventura, J. et Stamler, J.S. (1996) *Nature*, **380**, 221-226,
 - Kapp, O.H. et Crewe, A.V. (1984) *Biochim. Biophys. Acta* **789**, 294-301,

– Kapp, O.H., et al. (1984) The reassociation of *Lumbricus terrestris* hemoglobin dissociated at alkaline pH, *J Biol Chem*, **259**(1): 628-39,

– Krebs, A., et al. (1996) Molecular shape, dissociation, and oxygen binding of the dodecamer subunit of *Lumbricus terrestris* hemoglobin, *J Biol Chem*, **271**(31): 18695-704,

– Lamy, J.N.; Green, B.N.; Toulmond, A., Wall, J.S.; Weber, R.E. et Vinogradov, S.N. (1996), *Chem. Rev.* **96**, 3113-3124

– Levin, O. (1963) *J. Mol. Biol.*, **6**, 95-101,

– Mainwaring, M.G., et al. (1986) The dissociation of the extracellular hemoglobin of *Lumbricus terrestris* at acid pH and its reassociation at neutral pH. A new model of its quaternary structure, *J Biol Chem*, **261**(23): 10899-908,

– Mitsuno, T. et Naito, R. (1979) Perfluorochemical Blood Substitutes., Excerpta Medica, Amsterdam,

– Mitsuno, T. et Ohyanagi, H. (1985) Present status of clinical studies of fluosol-DA (20%) in Japan, K.K. Tremper (Ed), 169-184, Little Brown & Co, Boston,

– Naito, R. et Y. K. (1978) An improved perfluorodecalin emulsion. In Blood Substitutes and Plasma Expanders, G.A. Jamieson and T.J. Greenwalt (Eds), 81, Alan R. Liss Inc., New York,

– Nho, K.; Glower, D.; Bredehoeft, S.; Shankar, H.; Shorr, R. et Abuchowski, A. (1992) Biomaterials, Artificial Cells and Immobilization Biotechnology, An International Journal, **20**, 511-524,

– Payne, J.W. (1973) *Biochem J.* **135**, 866-873,

– Polidori, G., et al. (1984) The dissociation of the extracellular hemoglobin of *Tubifex tubifex* at extremes of pH and its reassociation upon return to neutrality, *Arch Biochem Biophys*, **233**(2): 800-14,

– Reiss, J.G. (1991) *Vox Sang*, **61**, 225-239,

– Reiss, J.G., (1994) Artificial Cells, Blood substitutes & Immobilization Biotechnology, An International Journal **22**, 945-1511,

– Roche, J. (1965) Electron microscope studies on high molecular weight erythrocrucorins (invertebrate hemoglobins) and chlorocruorins of annelids., D.A. Munday (Ed), 62-80, Pergamon Press, Oxford,

– Roche, J.; Bessis, M. et Thiery, J.P. (1960) *Biochim. Biophys. Acta*, **41**, 182-184,

– Roche, J.; Bessis, M.T. et Thiery, J.P. (1960) *C. R. Soc. Biol.* **154**, 73-80,

– Sharma, P.K., et al. (1996) The role of the dodecamer subunit in the dissociation and reassembly of the hexagonal bilayer structure of *Lumbricus terrestris* hemoglobin, *J Biol Chem*, **271**(15):8754-62;

– Sloviter, H. et Kamimoto, T. (1967) *Nature*, **216**, 458,

– Terwilliger, N.B. (1992) Molecular Structure of the extracellular heme proteins., Vol. 13, C.P. Mangum (Ed), 193-229, Springer-Verlag, Berlin,

– Van Bruggen, E.F.J. et Weber, R.E. (1974) *Biochim. Biophys. Acta*, **359**, 210-212,

– Vinogradov, S.N., et al. (1991) A dodecamer of globin chains is the principal functional subunit of the extracellular hemoglobin of *Lumbricus terrestris*, *J Biol Chem*, **266**(20): 13091-6,

– Vinogradov, S.N., Shlom J. M. et Doyle, M. (1979) Dissociation of the extracellular hemoglobin of *Arenicola cristata*. *Comp. Biochem. Physiol.*, **65B**: 145-150,

– Vinogradov, S.N. (1985) The structure of invertebrate extracellular hemoglobins (erythrocruorins and chlorocruorins), *Comp Biochem Physiol B*, **82**(1): 1-15,

– Zal, F.; Green, B.N.; Lallier, F.H.; Vinogradov, S.N. et Toulmond, A. (1997) *Eur. J. Biochem.*, **243**, 85-92.

REVENDICATIONS

1. Procédé de dissociation de la molécule d'hémoglobine extracellulaire d'Annélides, notamment d'*Arenicola marina*, permettant l'obtention des dix chaînes protéiques constituant ladite molécule,

ledit procédé étant caractérisé en ce qu'il comprend une étape de mise en présence d'un échantillon d'hémoglobine extracellulaire d'Annélides, notamment d'*Arenicola marina* avec au moins un agent dissociant, notamment un mélange constitué de dithiothréitol (DTT) ou de chlorhydrate de tris(2-carboxyéthyl)phosphine (TCEP) ou de bêta-mercaptoéthanol et d'un tampon de dissociation, pendant un temps suffisant pour séparer les chaînes protéiques les unes des autres.

2. Procédé de dissociation selon la revendication 1, caractérisé en ce que le tampon de dissociation comprend 0,1 M de base Trisma (tris[hydroxyméthyl]aminométhane) et 0 à 10 mM d'EDTA ajusté à un pH compris d'environ 5 à environ 12.

3. Procédé de dissociation selon la revendication 1 ou 2, caractérisé en ce que les dix chaînes protéiques constituant ladite molécule sont obtenues par la dissociation de quatre sous-unités par exemple en présence de DTT, lesdites sous-unités étant elles-mêmes obtenues par mise en présence d'un échantillon d'hémoglobine extracellulaire d'*Arenicola marina* avec différents agents dissociants, notamment un tampon de dissociation.

4. Procédé de préparation de couples d'amorces à partir des chaînes protéiques telles qu'obtenues selon le procédé tel que défini dans l'une des revendications 1 à 3, ledit procédé étant caractérisé en ce qu'il comprend les étapes suivantes :

– l'isolement de chacune des dix chaînes protéiques constituant la molécule d'hémoglobine telles qu'obtenues selon le procédé tel que défini selon l'une des revendications 1 à 3,

– le microséquençage de chacune des dix chaînes protéiques isolées susmentionnées par spectrométrie de masse, afin d'obtenir une microséquence correspondant à chacune des dix séquences constituées de 5 à 20 acides aminés, et

— la détermination de dix couples d'amorces dégénérées à partir des microséquences susmentionnées.

5 5. Couples d'amorces tels qu'obtenus selon le procédé de la revendication 4, lesdits couples étant notamment les suivants :

a) *Amorce sens* : ~~GAR TGY GGN CCN TTR CAR CG~~

Amorce antisens : CCA NGC NTC YTT RTC RAA GCA

10 b) *Amorce sens* : AN TGY GGN CCN CTN CAR CG

Amorce antisens : CCA NGC NTC YTT RTC RAA GCA

c) *Amorce sens* : AAR GTI AAR CAN AAC TGG

Amorce antisens : CCA NGC NCC DAT RTC RAA

15

d) *Amorce sens* : TGY TGY AGY ATH GAR GAY CG

Amorce antisens : CA NGC NYC RCT RTT RAA RCA

e) *Amorce sens* : GAY GAY TGY GTN TGY CC

20 *Amorce antisens* : ANA CRA ART CRT CNC CRC A

où :

R représente A ou G,

Y représente C ou T,

25 N représente A, G, C ou T,

I représente l'inosine,

D représente A, G ou T,

H représente A, C ou T

30 6. Procédé de préparation de séquences nucléotidiques codant pour les chaînes protéiques constituant la molécule d'hémoglobine d'*Arenicola marina*, à partir des amorces telles qu'obtenues selon le procédé de la revendication 4 ou 5, ledit procédé étant caractérisé en ce qu'il correspond à un procédé d'amplification en chaîne par

— la détermination de dix couples d'amorces dégénérées à partir des microséquences susmentionnées.

5 5. Couples d'amorces tels qu'obtenus selon le procédé de la revendication 4, lesdits couples étant notamment les suivants :

a) *Amorce sens* : GAR TGY GGN CCN TTR CAR CG
Amorce antisens : CCA NGC NTC YTT RTC RAA GCA

10 b) *Amorce sens* : AN TGY GGN CCN CTN CAR CG
Amorce antisens : CCA NGC NTC YTT RTC RAA GCA

c) *Amorce sens* : AAR GTI AAR CAN AAC TGG
Amorce antisens : CCA NGC NCC DAT RTC RAA

15 d) *Amorce sens* : TGY TGY AGY ATH GAR GAY CG
Amorce antisens : CA NGC NYC RCT RTT RAA RCA

e) *Amorce sens* : GAY GAY TGY GTN TGY CC
20 *Amorce antisens* : ANA CRA ART CRT CNC CRC A

où :

R représente A ou G,

Y représente C ou T,

25 N représente A, G, C ou T,

I représente l'inosine,

D représente A, G ou T,

H représente A, C ou T

30 6. Procédé de préparation de séquences nucléotidiques codant pour les chaînes protéiques constituant la molécule d'hémoglobine d'*Arenicola marina*, à partir des amorces telles qu'obtenues selon le procédé de la revendication 4, ledit procédé étant caractérisé en ce qu'il correspond à un procédé d'amplification en chaîne par

polymérase (PCR), comprenant une répétition d'au moins 30 fois du cycle constitué par les étapes suivantes :

* la dénaturation, par chauffage, de l'ADNc codant pour chacune des dix chaînes protéiques constituant la molécule d'hémoglobine d'*Arenicola marina*, de façon à séparer l'ADNc en deux brins monocaténares, ledit ADNc étant obtenu à partir d'ARNm mature,

* l'hybridation des couples d'amorces tels qu'obtenus par le procédé de la revendication 4 aux brins d'ADNc monocaténares susmentionnés à une température adéquate, pour obtenir des amorces hybridées, et

* l'élongation des amorces hybridées par une polymérase à une température adéquate.

7. Procédé de préparation selon la revendication 6, caractérisé en ce que :

- la première étape dudit procédé est une étape de dénaturation d'environ 10 secondes à environ 5 minutes à une température comprise d'environ 90°C à environ 110°C,

- le cycle, répété environ 30 à 40 fois, comprend les étapes suivantes :

* une étape de dénaturation d'environ 10 secondes à environ 5 minutes, à une température comprise d'environ 90°C à environ 110°C,

* une étape d'hybridation d'environ 20 secondes à environ 2 minutes, à une température comprise d'environ 50°C à environ 56°C,

* une étape d'élongation d'environ 20 secondes à environ 1 minute et 30 secondes, à une température comprise d'environ 70°C à environ 75°C, et

- la dernière étape du procédé est une étape d'élongation d'environ 5 minutes à environ 15 minutes à une température comprise d'environ 70°C à environ 75°C.

8. Procédé de préparation selon la revendication 6 ou 7, caractérisé en ce que les couples d'amorces utilisés sont tels que définis dans la revendication 5.

9. Procédé de préparation selon la revendication 8, caractérisé en ce que le couple d'amorces utilisé est le suivant : (GAR TGY GGN CCN TTR CAR CG ; CCA NGC NTC YTT RTC RAA GCA), R, Y et N étant tels que définis dans la revendication 5, ledit procédé étant caractérisé en ce que :

– la première étape du procédé est une étape de dénaturation de 4 minutes à une température égale à 95°C,

– le cycle, répété 35 fois, comprend les étapes suivantes :

* une étape de dénaturation de 30 secondes à une température égale à 95°C,

* une étape d'hybridation de 30 secondes à une température égale à 56°C,

* une étape d'élongation de 40 secondes à une température égale à 72°C, et

– la dernière étape du procédé est une étape d'élongation de 10 minutes à une température égale à 72°C,

afin d'obtenir la séquence nucléotidique SEQ ID NO : 1.

10. Procédé de préparation selon la revendication 8, caractérisé en ce que le couple d'amorces utilisé est le suivant : (AN TGY GGN CCN CTN CAR CG ; CCA NGC NTC YTT RTC RAA GCA), N, Y et R étant tels que définis dans la revendication 5, ledit procédé étant caractérisé en ce que :

– la première étape du procédé est une étape de dénaturation de 4 minutes à une température égale à 95°C,

– le cycle, répété 35 fois, comprend les étapes suivantes :

* une étape de dénaturation de 30 secondes à une température égale à 95°C,

* une étape d'hybridation de 30 secondes à une température égale à 52°C,

* une étape d'élongation de 40 secondes à une température égale à 72°C, et

– la dernière étape du procédé est une étape d'élongation de 10 minutes à une température égale à 72°C,

afin d'obtenir la séquence nucléotidique SEQ ID NO : 3.

11. Procédé de préparation selon la revendication 8, caractérisé en ce que le couple d'amorces utilisé est le suivant : (AAR GTI AAR CAN AAC TGG ; CCA NGC NCC DAT RTC RAA), R, I, N et D étant tels que définis dans la revendication 5, ledit procédé étant caractérisé en ce que :

– la première étape du procédé est une étape de dénaturation de 4 minutes à une température égale à 95°C,

– le cycle, répété 35 fois, comprend les étapes suivantes :

* une étape de dénaturation de 1 minute à une température égale à 95°C,

* une étape d'hybridation de 1 minute à une température égale à 50°C,

- * une étape d'élongation de 1 minute et 30 secondes à une température égale à 72°C, et

— la dernière étape du procédé est une étape d'élongation de 10 minutes à une température égale à 72°C,

afin d'obtenir la séquence nucléotidique SEQ ID NO : 5.

12. Procédé de préparation selon la revendication 8, caractérisé en ce que le couple d'amorces utilisé est le suivant : (TGY TGY AGY ATH GAR GAY CG ; CA NGC NYC RCT RTT RAA RCA), Y, H, R et N étant tels que définis dans la revendication 5,

ledit procédé étant caractérisé en ce que :

- la première étape du procédé est une étape de dénaturation de 4 minutes à une température égale à 95°C,

- le cycle, répété 35 fois, comprend les étapes suivantes :

- * une étape de dénaturation de 30 secondes à une température égale à 95°C,
- * une étape d'hybridation de 40 secondes à une température égale à 52°C,
- * une étape d'élongation de 30 secondes à une température égale à 72°C, et

- la dernière étape du procédé est une étape d'élongation de 10 minutes à une température égale à 72°C,

afin d'obtenir la séquence nucléotidique SEQ ID NO : 7.

13. Protéines codées par l'une des séquences telles qu'obtenues selon le procédé tel que défini dans l'une quelconque des revendications 6 à 12.

14. Protéine selon la revendication 13, caractérisée en ce qu'elle comprend ou est constituée par :

- la séquence SEQ ID NO : 2,

— ou toute séquence dérivée de la séquence SEQ ID NO : 2 ou d'un fragment défini ci-dessous, notamment par substitution, suppression ou addition d'un ou plusieurs acides aminés, sous réserve que ladite séquence dérivée permette le transport de l'oxygène,

— ou toute séquence homologue de la séquence SEQ ID NO : 2 ou d'un fragment défini ci-dessous, ayant de préférence une homologie d'au moins environ 75%, avec la

séquence SEQ ID NO : 2, sous réserve que ladite séquence homologue permette le transport de l'oxygène,

– ou tout fragment d'une des séquences définies ci-dessus, sous réserve que ledit fragment permette le transport de l'oxygène, notamment tout fragment étant constitué d'au moins environ 60 acides aminés, et notamment d'au moins environ 160 acides aminés contigus dans la séquence SEQ ID NO : 2.

15. Protéine selon la revendication 13, caractérisée en ce qu'elle comprend ou est constituée par :

– la séquence SEQ ID NO : 4,

– ou toute séquence dérivée de la séquence SEQ ID NO : 4 ou d'un fragment défini ci-dessous, notamment par substitution, suppression ou addition d'un ou plusieurs acides aminés, sous réserve que ladite séquence dérivée permette le transport de l'oxygène,

– ou toute séquence homologue de la séquence SEQ ID NO : 4 ou d'un fragment défini ci-dessous, ayant de préférence une homologie d'au moins environ 75%, avec la séquence SEQ ID NO : 4, sous réserve que ladite séquence homologue permette le transport de l'oxygène,

– ou tout fragment d'une des séquences définies ci-dessus, sous réserve que ledit fragment permette le transport de l'oxygène, notamment tout fragment étant constitué d'au moins environ 60 acides aminés, et notamment d'au moins environ 160 acides aminés contigus dans la séquence SEQ ID NO : 4.

16. Protéine selon la revendication 13, caractérisée en ce qu'elle comprend ou est constituée par :

– la séquence SEQ ID NO : 6,

– ou toute séquence dérivée de la séquence SEQ ID NO : 6 ou d'un fragment défini ci-dessous, notamment par substitution, suppression ou addition d'un ou plusieurs acides aminés, sous réserve que ladite séquence dérivée permette le transport de l'oxygène,

– ou toute séquence homologue de la séquence SEQ ID NO : 6 ou d'un fragment défini ci-dessous, ayant de préférence une homologie d'au moins environ 75%, avec la séquence SEQ ID NO : 6, sous réserve que ladite séquence homologue permette le transport de l'oxygène,

séquence SEQ ID NO : 2, sous réserve que ladite séquence homologue permette le transport de l'oxygène,

— ou tout fragment d'une des séquences définies ci-dessus, sous réserve que ledit fragment permette le transport de l'oxygène, notamment tout fragment étant constitué d'au moins environ 60 acides aminés contigus dans la séquence SEQ ID NO : 2.

15. Protéine selon la revendication 13, caractérisée en ce qu'elle comprend ou est constituée par :

— la séquence SEQ ID NO : 4,

— ou toute séquence dérivée de la séquence SEQ ID NO : 4 ou d'un fragment défini ci-dessous, notamment par substitution, suppression ou addition d'un ou plusieurs acides aminés, sous réserve sous réserve que ladite séquence dérivée permette le transport de l'oxygène,

— ou toute séquence homologue de la séquence SEQ ID NO : 4 ou d'un fragment défini ci-dessous, ayant de préférence une homologie d'au moins environ 75%, avec la séquence SEQ ID NO : 4, sous réserve que ladite séquence homologue permette le transport de l'oxygène,

— ou tout fragment d'une des séquences définies ci-dessus, sous réserve que ledit fragment permette le transport de l'oxygène, notamment tout fragment étant constitué d'au moins environ 60 acides aminés contigus dans la séquence SEQ ID NO : 4.

16. Protéine selon la revendication 13, caractérisée en ce qu'elle comprend ou est constituée par :

— la séquence SEQ ID NO : 6,

— ou toute séquence dérivée de la séquence SEQ ID NO : 6 ou d'un fragment défini ci-dessous, notamment par substitution, suppression ou addition d'un ou plusieurs acides aminés, sous réserve sous réserve que ladite séquence dérivée permette le transport de l'oxygène,

— ou toute séquence homologue de la séquence SEQ ID NO : 6 ou d'un fragment défini ci-dessous, ayant de préférence une homologie d'au moins environ 75%, avec la séquence SEQ ID NO : 6, sous réserve que ladite séquence homologue permette le transport de l'oxygène,

– ou tout fragment d'une des séquences définies ci-dessus, sous réserve que ledit fragment permette le transport de l'oxygène, notamment tout fragment étant constitué d'au moins environ 60 acides aminés, et notamment d'au moins environ 160 acides aminés contigus dans la séquence SEQ ID NO : 6.

5

17. Protéine selon la revendication 13, caractérisée en ce qu'elle comprend ou est constituée par :

10

– la séquence SEQ ID NO : 8,

– ou toute séquence dérivée de la séquence SEQ ID NO : 8 ou d'un fragment défini ci-dessous, notamment par substitution, suppression ou addition d'un ou plusieurs acides aminés, sous réserve que ladite séquence dérivée permette le transport de l'oxygène,

15

– ou toute séquence homologue de la séquence SEQ ID NO : 8 ou d'un fragment défini ci-dessous, ayant de préférence une homologie d'au moins environ 75%, avec la séquence SEQ ID NO : 8, sous réserve que ladite séquence homologue permette le transport de l'oxygène,

20

– ou tout fragment d'une des séquences définies ci-dessus, sous réserve que ledit fragment permette le transport de l'oxygène, notamment tout fragment étant constitué d'au moins environ 60 acides aminés, et notamment d'au moins environ 160 acides aminés contigus dans la séquence SEQ ID NO : 8.

18. Séquences nucléotidiques telles qu'obtenues selon le procédé tel que défini dans l'une quelconque des revendications 6 à 12.

25

19. Séquences nucléotidiques codant pour une protéine telle que définie dans l'une des revendications 13 à 17.

20. Séquence nucléotidique selon la revendication 19, caractérisée en ce qu'elle comprend ou est constituée par :

30

– la séquence nucléotidique SEQ ID NO : 1 codant pour SEQ ID NO : 2,

– ou toute séquence nucléotidique dérivée, par dégénérescence du code génétique, de la séquence SEQ ID NO : 1, et codant pour une protéine représentée par SEQ ID NO : 2,

— ou tout fragment d'une des séquences définies ci-dessus, sous réserve que ledit fragment permette le transport de l'oxygène, notamment tout fragment étant constitué d'au moins environ 60 acides aminés contigus dans la séquence SEQ ID NO : 6.

5 17. Protéine selon la revendication 13, caractérisée en ce qu'elle comprend ou est constituée par :

— la séquence SEQ ID NO : 8,
— ou toute séquence dérivée de la séquence SEQ ID NO : 8 ou d'un fragment défini ci-dessous, notamment par substitution, suppression ou addition d'un ou plusieurs
10 acides aminés, sous réserve que ladite séquence dérivée permette le transport de l'oxygène,

— ou toute séquence homologue de la séquence SEQ ID NO : 8 ou d'un fragment défini ci-dessous, ayant de préférence une homologie d'au moins environ 75%, avec la séquence SEQ ID NO : 8, sous réserve que ladite séquence homologue permette le
15 transport de l'oxygène,

— ou tout fragment d'une des séquences définies ci-dessus, sous réserve que ledit fragment permette le transport de l'oxygène, notamment tout fragment étant constitué d'au moins environ 60 acides aminés contigus dans la séquence SEQ ID NO : 8.

20 18. Séquences nucléotidiques telles qu'obtenues selon le procédé tel que défini dans l'une quelconque des revendications 6 à 12.

19. Séquences nucléotidiques codant pour une protéine telle que définie dans l'une des revendications 13 à 17.

25 20. Séquence nucléotidique selon la revendication 19, caractérisée en ce qu'elle comprend ou est constituée par :

— la séquence nucléotidique SEQ ID NO : 1 codant pour SEQ ID NO : 2,
— ou toute séquence nucléotidique dérivée, par dégénérescence du code
30 génétique, de la séquence SEQ ID NO : 1, et codant pour une protéine représentée par SEQ ID NO : 2,

— ou toute séquence nucléotidique dérivée, notamment par substitution, suppression ou addition d'un ou plusieurs nucléotides, de la séquence SEQ ID NO : 1 codant pour une protéine dérivée de SEQ ID NO : 2,

– ou toute séquence nucléotidique dérivée, notamment par substitution, suppression ou addition d'un ou plusieurs nucléotides, de la séquence SEQ ID NO : 1 codant pour une protéine dérivée de SEQ ID NO : 2,

– ou toute séquence nucléotidique homologue de SEQ ID NO : 1, ayant de préférence une homologie d'au moins environ 60%, avec la séquence SEQ ID NO : 1,

– ou tout fragment de la séquence nucléotidique SEQ ID NO : 1 ou des séquences nucléotidiques définies ci-dessus, ledit fragment étant de préférence constitué d'au moins environ 180 nucléotides, et notamment d'au moins environ 480 nucléotides contigus dans ladite séquence,

– ou toute séquence nucléotidique complémentaire des séquences ou fragments susmentionnés,

– ou toute séquence nucléotidique susceptible d'hybrider dans des conditions stringentes avec la séquence complémentaire d'une des séquences ou d'un des fragments susmentionnés.

21. Séquence nucléotidique selon la revendication 19, caractérisée en ce qu'elle comprend ou est constituée par :

– la séquence nucléotidique SEQ ID NO : 3 codant pour SEQ ID NO : 4,

– ou toute séquence nucléotidique dérivée, par dégénérescence du code génétique, de la séquence SEQ ID NO : 3, et codant pour une protéine représentée par SEQ ID NO : 4,

– ou toute séquence nucléotidique dérivée, notamment par substitution, suppression ou addition d'un ou plusieurs nucléotides, de la séquence SEQ ID NO : 3 codant pour une protéine dérivée de SEQ ID NO : 4,

– ou toute séquence nucléotidique homologue de SEQ ID NO : 3, ayant de préférence une homologie d'au moins environ 60%, avec la séquence SEQ ID NO : 3,

– ou tout fragment de la séquence nucléotidique SEQ ID NO : 3 ou des séquences nucléotidiques définies ci-dessus, ledit fragment étant de préférence constitué d'au moins environ 180 nucléotides, et notamment d'au moins environ 480 nucléotides contigus dans ladite séquence,

– ou toute séquence nucléotidique complémentaire des séquences ou fragments susmentionnés,

– ou toute séquence nucléotidique homologue de SEQ ID NO : 1, ayant de préférence une homologie d'au moins environ 60%, avec la séquence SEQ ID NO : 1,

– ou tout fragment de la séquence nucléotidique SEQ ID NO : 1 ou des séquences nucléotidiques définies ci-dessus, ledit fragment étant de préférence constitué d'au moins environ 180 nucléotides contigus dans ladite séquence,

– ou toute séquence nucléotidique complémentaire des séquences ou fragments susmentionnés,

– ou toute séquence nucléotidique susceptible d'hybrider dans des conditions stringentes avec la séquence complémentaire d'une des séquences ou d'un des fragments susmentionnés.

21. Séquence nucléotidique selon la revendication 19, caractérisée en ce qu'elle comprend ou est constituée par :

– la séquence nucléotidique SEQ ID NO : 3 codant pour SEQ ID NO : 4,

– ou toute séquence nucléotidique dérivée, par dégénérescence du code génétique, de la séquence SEQ ID NO : 3, et codant pour une protéine représentée par SEQ ID NO : 4,

– ou toute séquence nucléotidique dérivée, notamment par substitution, suppression ou addition d'un ou plusieurs nucléotides, de la séquence SEQ ID NO : 3 codant pour une protéine dérivée de SEQ ID NO : 4,

– ou toute séquence nucléotidique homologue de SEQ ID NO : 3, ayant de préférence une homologie d'au moins environ 60%, avec la séquence SEQ ID NO : 3,

– ou tout fragment de la séquence nucléotidique SEQ ID NO : 3 ou des séquences nucléotidiques définies ci-dessus, ledit fragment étant de préférence constitué d'au moins environ 180 nucléotides contigus dans ladite séquence,

– ou toute séquence nucléotidique complémentaire des séquences ou fragments susmentionnés,

– ou toute séquence nucléotidique susceptible d'hybrider dans des conditions stringentes avec la séquence complémentaire d'une des séquences ou d'un des fragments susmentionnés.

— ou toute séquence nucléotidique susceptible d'hybrider dans des conditions stringentes avec la séquence complémentaire d'une des séquences ou d'un des fragments susmentionnés.

5 22. Séquence nucléotidique selon la revendication 19, caractérisée en ce qu'elle comprend ou est constituée par :

— la séquence nucléotidique SEQ ID NO : 5 codant pour SEQ ID NO : 6,

— ou toute séquence nucléotidique dérivée, par dégénérescence du code génétique, de la séquence SEQ ID NO : 5, et codant pour une protéine représentée par
10 SEQ ID NO : 6,

— ou toute séquence nucléotidique dérivée, notamment par substitution, suppression ou addition d'un ou plusieurs nucléotides, de la séquence SEQ ID NO : 5 codant pour une protéine dérivée de SEQ ID NO : 6,

— ou toute séquence nucléotidique homologue de SEQ ID NO : 5, ayant de
15 préférence une homologie d'au moins environ 60%, avec la séquence SEQ ID NO : 5,

— ou tout fragment de la séquence nucléotidique SEQ ID NO : 5 ou des séquences nucléotidiques définies ci-dessus, ledit fragment étant de préférence constitué d'au moins environ 180 nucléotides, et notamment d'au moins environ 480 nucléotides
contigus dans ladite séquence,

20 — ou toute séquence nucléotidique complémentaire des séquences ou fragments susmentionnés,

— ou toute séquence nucléotidique susceptible d'hybrider dans des conditions stringentes avec la séquence complémentaire d'une des séquences ou d'un des fragments susmentionnés.

25 23. Séquence nucléotidique selon la revendication 19, caractérisée en ce qu'elle comprend ou est constituée par :

— la séquence nucléotidique SEQ ID NO : 7 codant pour SEQ ID NO : 8,

— ou toute séquence nucléotidique dérivée, par dégénérescence du code
30 génétique, de la séquence SEQ ID NO : 7, et codant pour une protéine représentée par
SEQ ID NO : 8,

22. Séquence nucléotidique selon la revendication 19, caractérisée en ce qu'elle comprend ou est constituée par :

- la séquence nucléotidique SEQ ID NO : 5 codant pour SEQ ID NO : 6,
- ou toute séquence nucléotidique dérivée, par dégénérescence du code génétique, de la séquence SEQ ID NO : 5, et codant pour une protéine représentée par SEQ ID NO : 6,
- ou toute séquence nucléotidique dérivée, notamment par substitution, suppression ou addition d'un ou plusieurs nucléotides, de la séquence SEQ ID NO : 5 codant pour une protéine dérivée de SEQ ID NO : 6,
- ou toute séquence nucléotidique homologue de SEQ ID NO : 5, ayant de préférence une homologie d'au moins environ 60%, avec la séquence SEQ ID NO : 5,
- ou tout fragment de la séquence nucléotidique SEQ ID NO : 5 ou des séquences nucléotidiques définies ci-dessus, ledit fragment étant de préférence constitué d'au moins environ 180 nucléotides contigus dans ladite séquence,
- ou toute séquence nucléotidique complémentaire des séquences ou fragments susmentionnés,
- ou toute séquence nucléotidique susceptible d'hybrider dans des conditions stringentes avec la séquence complémentaire d'une des séquences ou d'un des fragments susmentionnés.

23. Séquence nucléotidique selon la revendication 19, caractérisée en ce qu'elle comprend ou est constituée par :

- la séquence nucléotidique SEQ ID NO : 7 codant pour SEQ ID NO : 8,
- ou toute séquence nucléotidique dérivée, par dégénérescence du code génétique, de la séquence SEQ ID NO : 7, et codant pour une protéine représentée par SEQ ID NO : 8,
- ou toute séquence nucléotidique dérivée, notamment par substitution, suppression ou addition d'un ou plusieurs nucléotides, de la séquence SEQ ID NO : 7 codant pour une protéine dérivée de SEQ ID NO : 8,
- ou toute séquence nucléotidique homologue de SEQ ID NO : 7, ayant de préférence une homologie d'au moins environ 60%, avec la séquence SEQ ID NO : 7,

– ou toute séquence nucléotidique dérivée, notamment par substitution, suppression ou addition d'un ou plusieurs nucléotides, de la séquence SEQ ID NO : 7 codant pour une protéine dérivée de SEQ ID NO : 8,

– ou toute séquence nucléotidique homologue de SEQ ID NO : 7, ayant de préférence une homologie d'au moins environ 60%, avec la séquence SEQ ID NO : 7,

– ou tout fragment de la séquence nucléotidique SEQ ID NO : 7 ou des séquences nucléotidiques définies ci-dessus, ledit fragment étant de préférence constitué d'au moins environ 180 nucléotides, et notamment d'au moins environ 480 nucléotides contigus dans ladite séquence,

– ou toute séquence nucléotidique complémentaire des séquences ou fragments susmentionnés,

– ou toute séquence nucléotidique susceptible d'hybrider dans des conditions stringentes avec la séquence complémentaire d'une des séquences ou d'un des fragments susmentionnés.

24. Procédé de préparation selon la revendication 4, de séquences nucléotidiques codant pour les chaînes protéiques constituant la molécule d'hémoglobine d'Annélides, notamment d'*Arenicola marina*, ledit procédé étant caractérisé en ce qu'il comprend les étapes suivantes :

– une étape de mise en présence de la molécule d'hémoglobine susmentionnée avec au moins un agent dissociant, notamment un mélange constitué de dithiothréitol (DTT) ou de chlorhydrate de tris(2-carboxyéthyl)phosphine (TCEP) ou de bêta-mercaptoéthanol et d'un tampon de dissociation, pendant un temps suffisant pour séparer les chaînes protéiques les unes des autres,

permettant la dissociation de ladite molécule d'hémoglobine, afin d'obtenir les dix chaînes protéiques constituant ladite molécule,

– l'isolement de chacune des dix chaînes protéiques susmentionnées,

– le microséquençage de chacune des dix chaînes protéiques isolées susmentionnées par spectrométrie de masse, afin d'obtenir une microséquence correspondant à chacune des dix séquences constituées de 5 à 20 acides aminés,

– la détermination des dix couples d'amorces dégénérées à partir des microséquences susmentionnées,

– ou tout fragment de la séquence nucléotidique SEQ ID NO : 7 ou des séquences nucléotidiques définies ci-dessus, ledit fragment étant de préférence constitué d'au moins environ 180 nucléotides contigus dans ladite séquence,

– ou toute séquence nucléotidique complémentaire des séquences ou fragments susmentionnés,

– ou toute séquence nucléotidique susceptible d'hybrider dans des conditions stringentes avec la séquence complémentaire d'une des séquences ou d'un des fragments susmentionnés.

24. Procédé de préparation selon la revendication 6, de séquences nucléotidiques codant pour les chaînes protéiques constituant la molécule d'hémoglobine d'Annélides, notamment d'*Arenicola marina*, ledit procédé étant caractérisé en ce qu'il comprend les étapes suivantes :

– une étape de mise en présence de la molécule d'hémoglobine susmentionnée avec au moins un agent dissolvant, notamment un mélange constitué de dithiothréitol (DTT) ou de chlorhydrate de tris(2-carboxyéthyl)phosphine (TCEP) ou de bêta-mercaptoéthanol et d'un tampon de dissociation, pendant un temps suffisant pour séparer les chaînes protéiques les unes des autres,

permettant la dissociation de ladite molécule d'hémoglobine, afin d'obtenir les dix chaînes protéiques constituant ladite molécule,

– l'isolement de chacune des dix chaînes protéiques susmentionnées,

– le microséquençage de chacune des dix chaînes protéiques isolées susmentionnées par spectrométrie de masse, afin d'obtenir une microséquence correspondant à chacune des dix séquences constituées de 5 à 20 acides aminés,

– la détermination des dix couples d'amorces dégénérées à partir des microséquences susmentionnées,

– la préparation des séquences nucléotidiques codant pour les chaînes protéiques susmentionnées, à partir des amorces telles qu'obtenues précédemment, par un procédé d'amplification en chaîne par polymérase (PCR), comprenant les étapes suivantes :

– la première étape dudit procédé est une étape de dénaturation d'environ 10 secondes à environ 5 minutes à une température comprise d'environ 90°C à environ 110°C,

– la préparation des séquences nucléotidiques codant pour les chaînes protéiques susmentionnées, à partir des amorces telles qu'obtenues précédemment, par un procédé d'amplification en chaîne par polymérase (PCR), comprenant les étapes suivantes :

– la première étape dudit procédé est une étape de dénaturation d'environ 10 secondes à environ 5 minutes à une température comprise d'environ 90°C à environ 110°C,

– le cycle, répété environ 30 à 40 fois, comprend les étapes suivantes :

* une étape de dénaturation d'environ 10 secondes à environ 5 minutes, à une température comprise d'environ 90°C à environ 110°C,

* une étape d'hybridation d'environ 20 secondes à environ 2 minutes, à une température comprise d'environ 50°C à environ 56°C,

* une étape d'élongation d'environ 20 secondes à environ 1 minute et 30 secondes, à une température comprise d'environ 70°C à environ 75°C, et

– la dernière étape du procédé est une étape d'élongation d'environ 5 minutes à environ 15 minutes à une température comprise d'environ 70°C à environ 75°C.

- le cycle, répété environ 30 à 40 fois, comprend les étapes suivantes :
 - * une étape de dénaturation d'environ 10 secondes à environ 5 minutes, à une température comprise d'environ 90°C à environ 110°C,
 - * une étape d'hybridation d'environ 20 secondes à environ 2 minutes, à une température comprise d'environ 50°C à environ 56°C,
 - * une étape d'élongation d'environ 20 secondes à environ 1 minute et 30 secondes, à une température comprise d'environ 70°C à environ 75°C, et
- la dernière étape du procédé est une étape d'élongation d'environ 5 minutes à environ 15 minutes à une température comprise d'environ 70°C à environ 75°C.

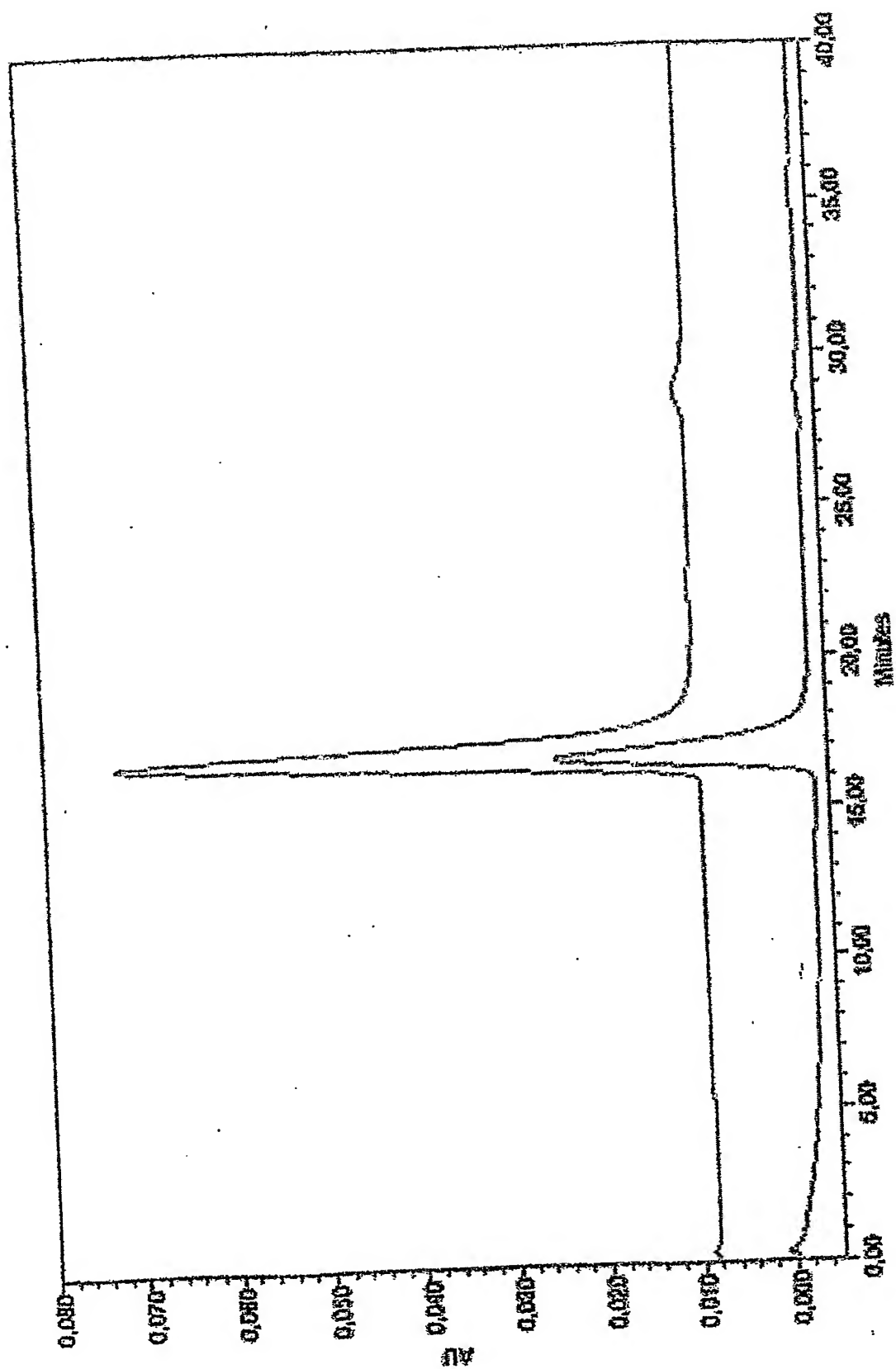


FIGURE 1

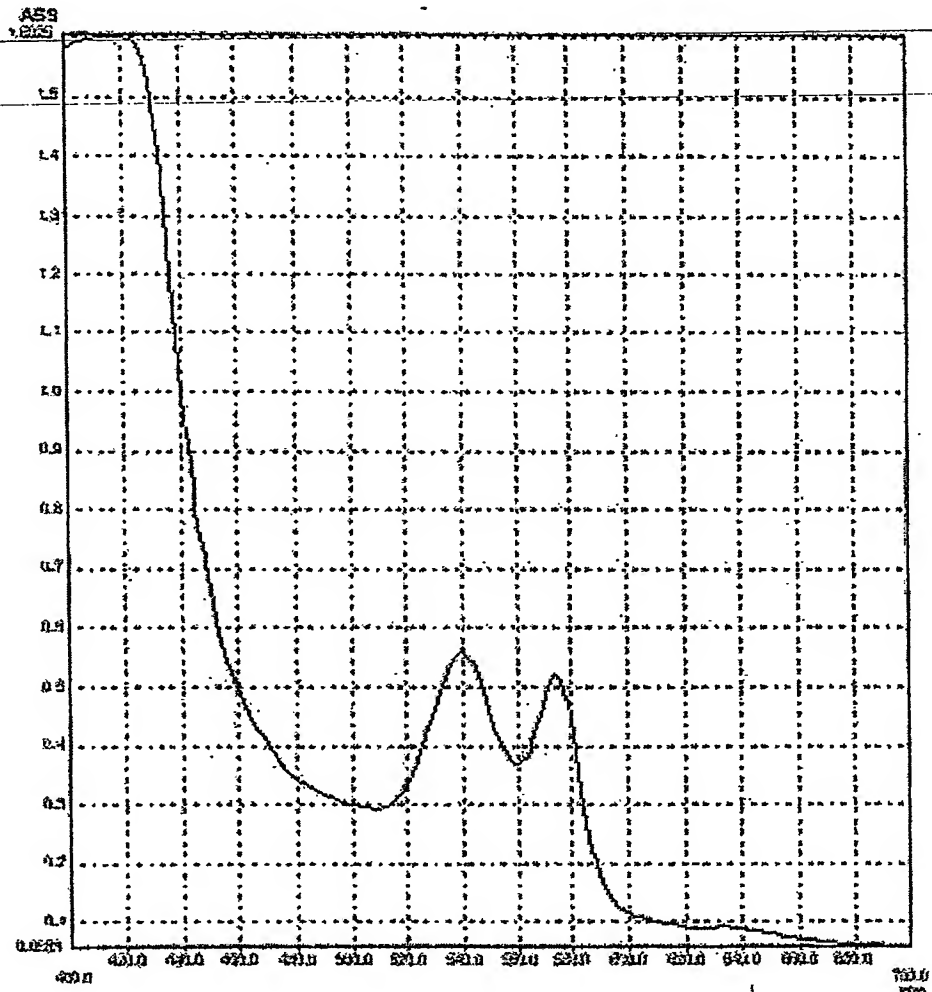


FIGURE 2

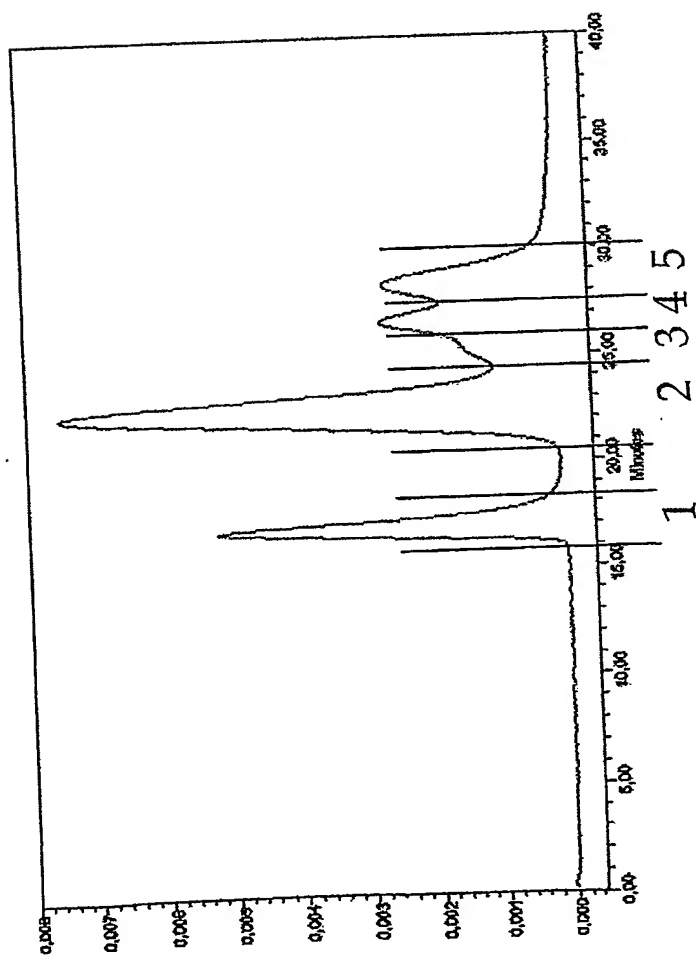


FIGURE 3

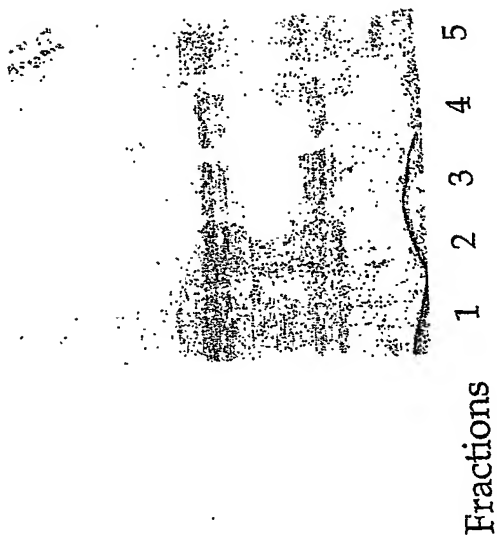


FIGURE 4

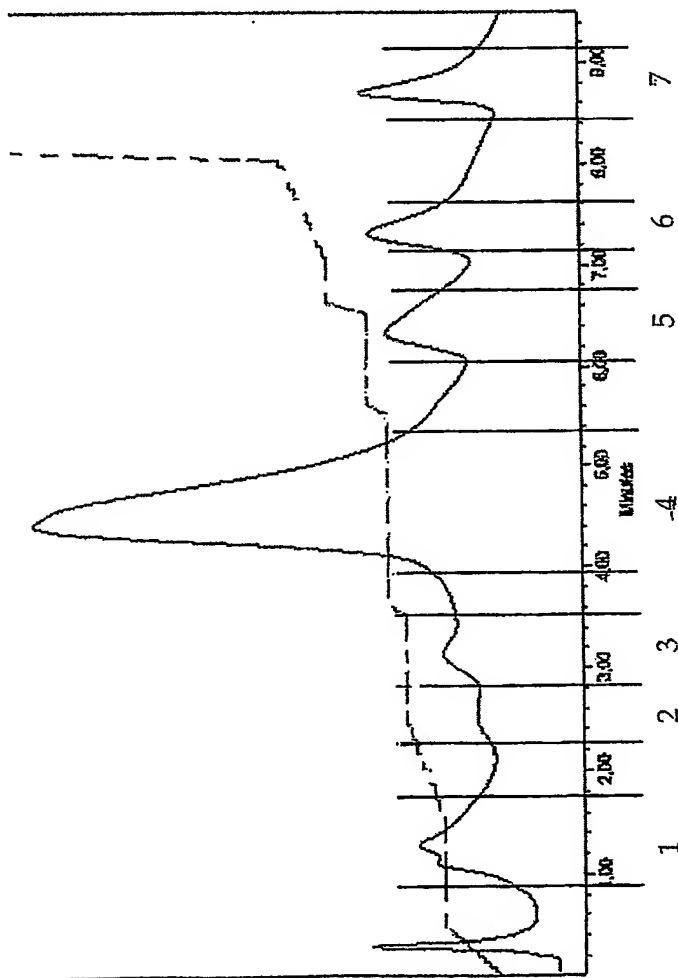


FIGURE 5

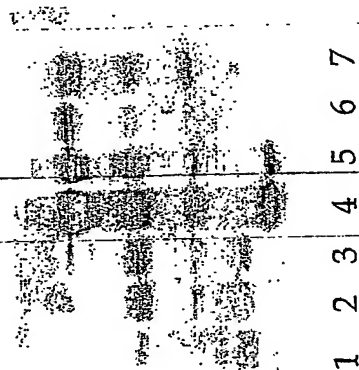


FIGURE 6

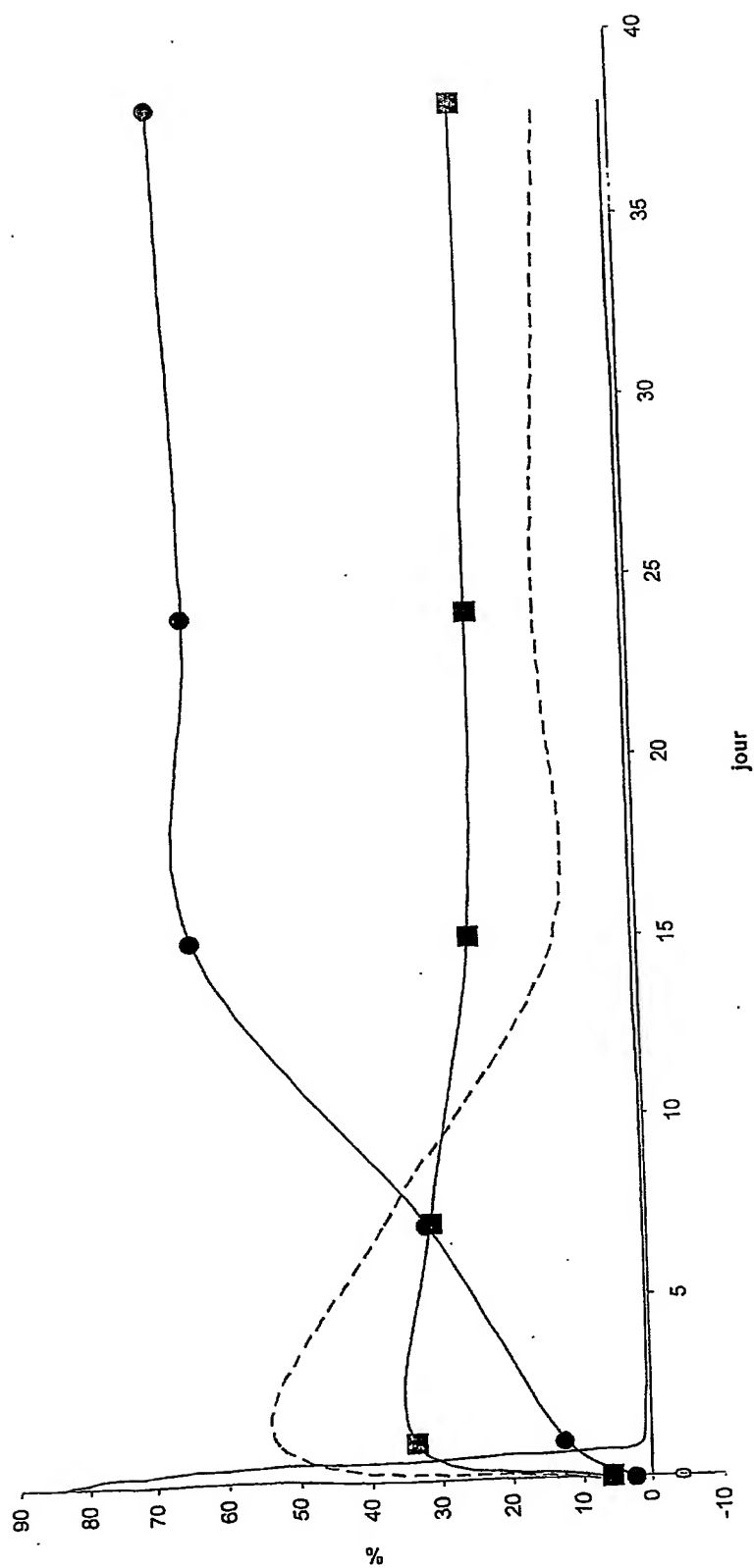


FIGURE 7

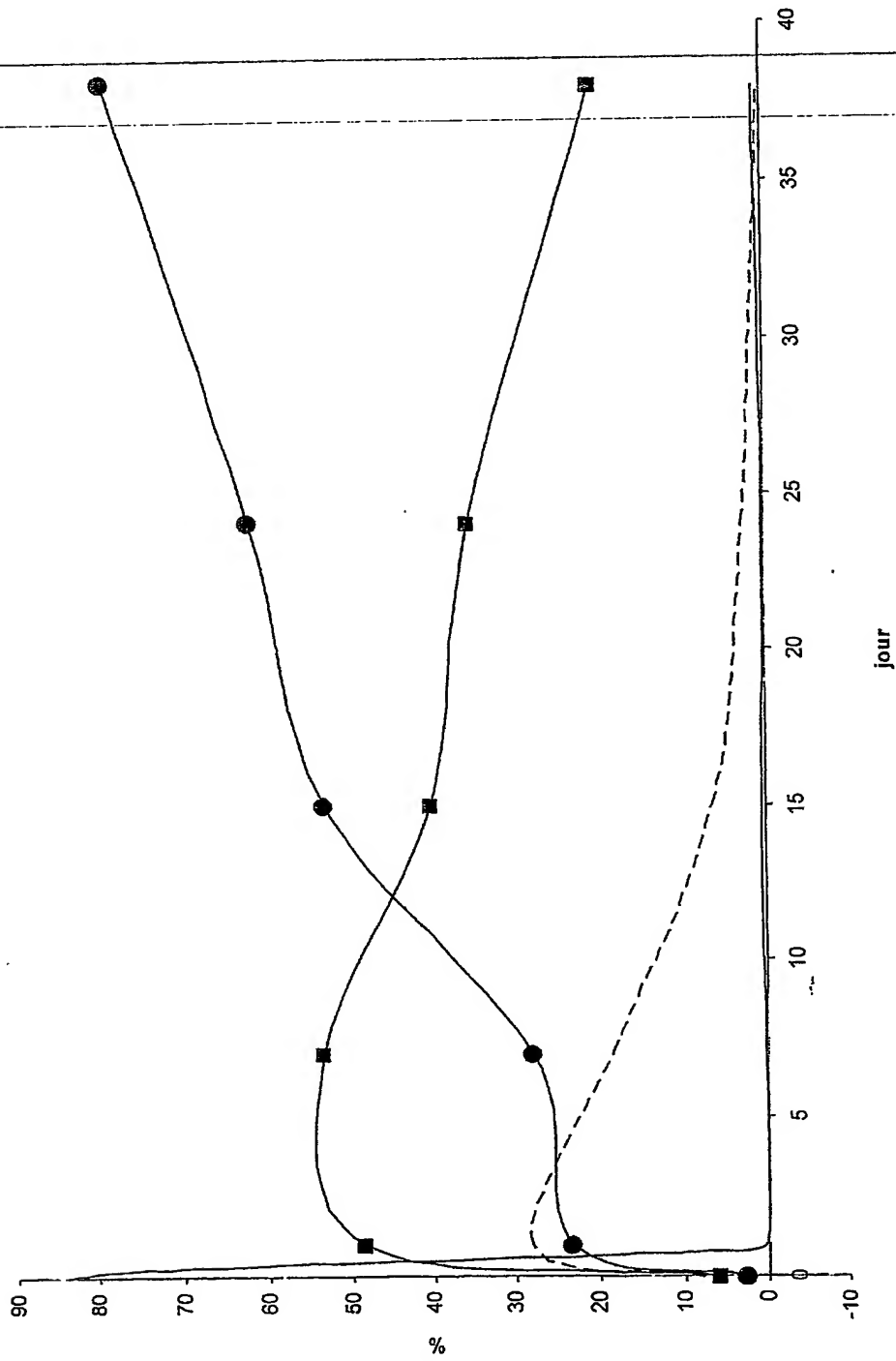


FIGURE 8

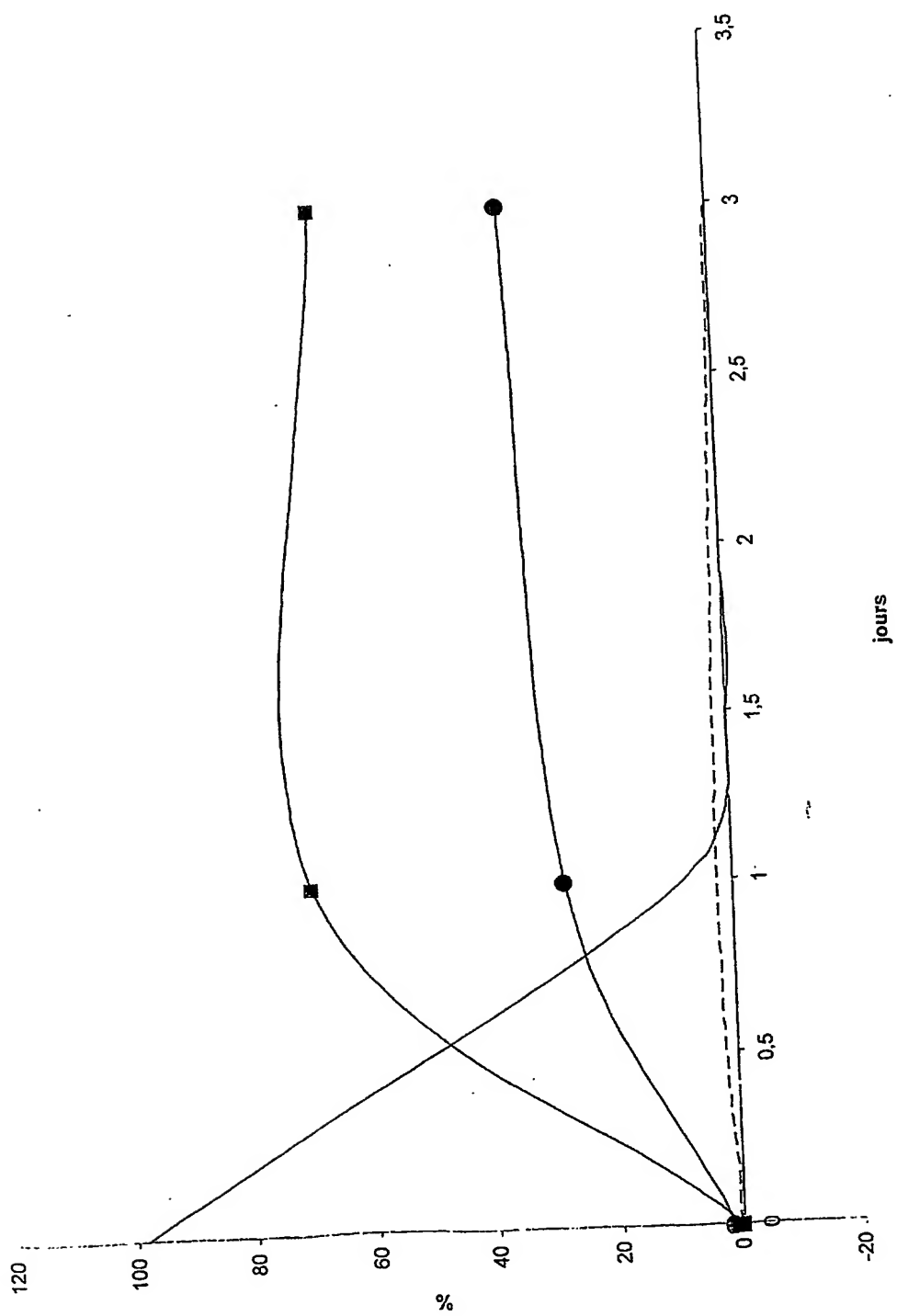


FIGURE 9

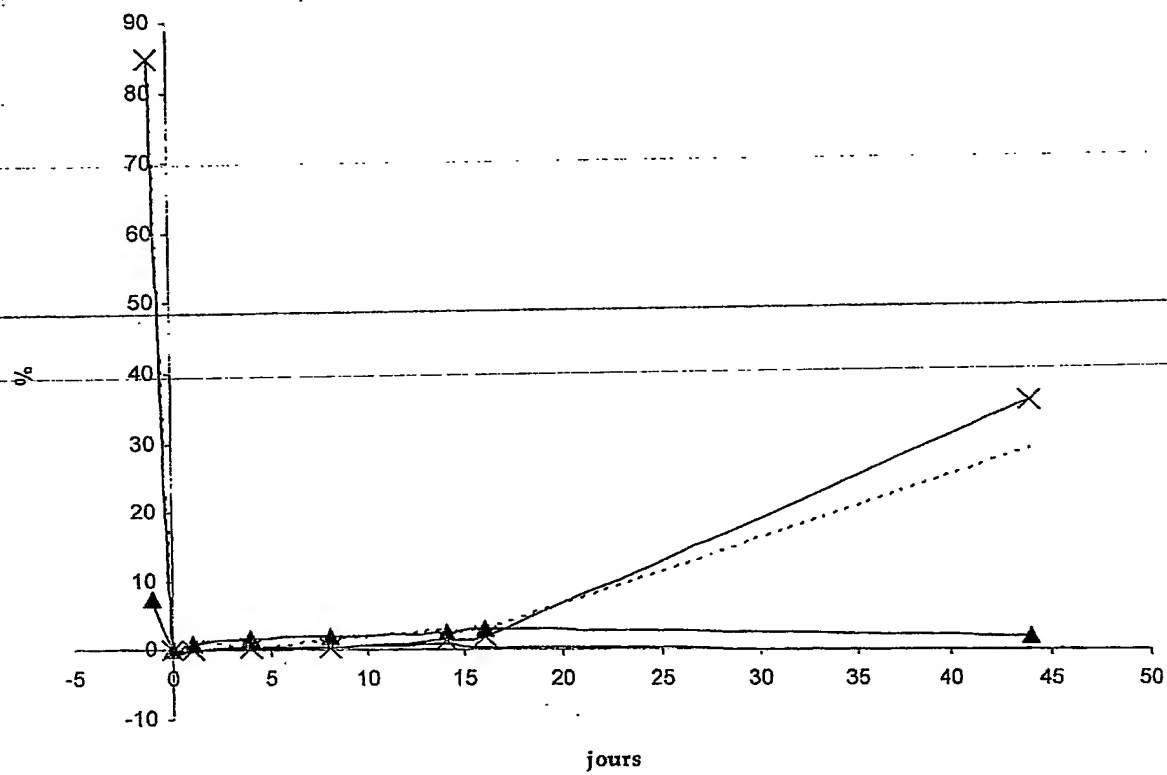


FIGURE 10

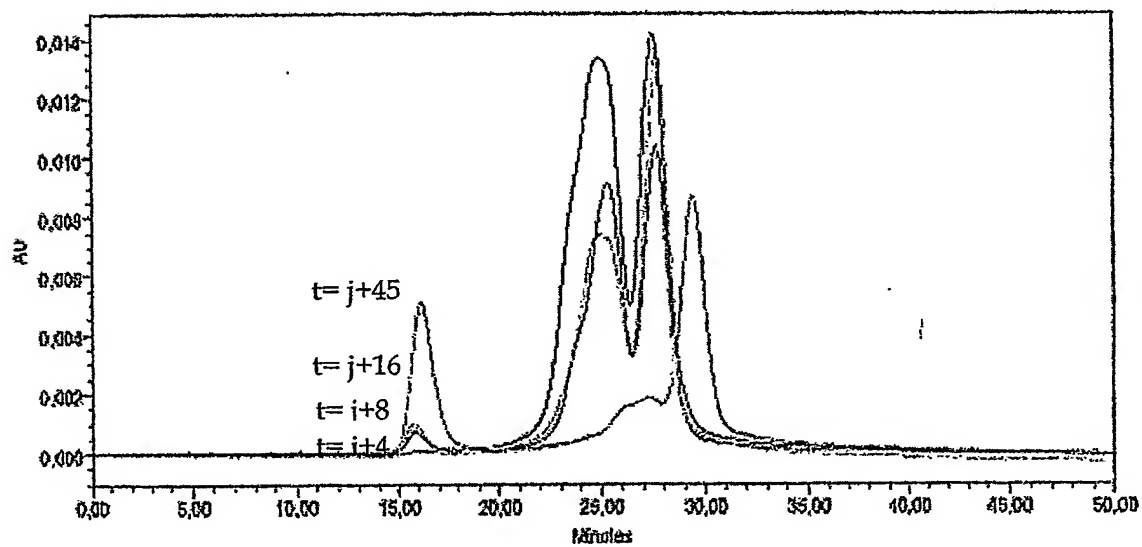


FIGURE 11

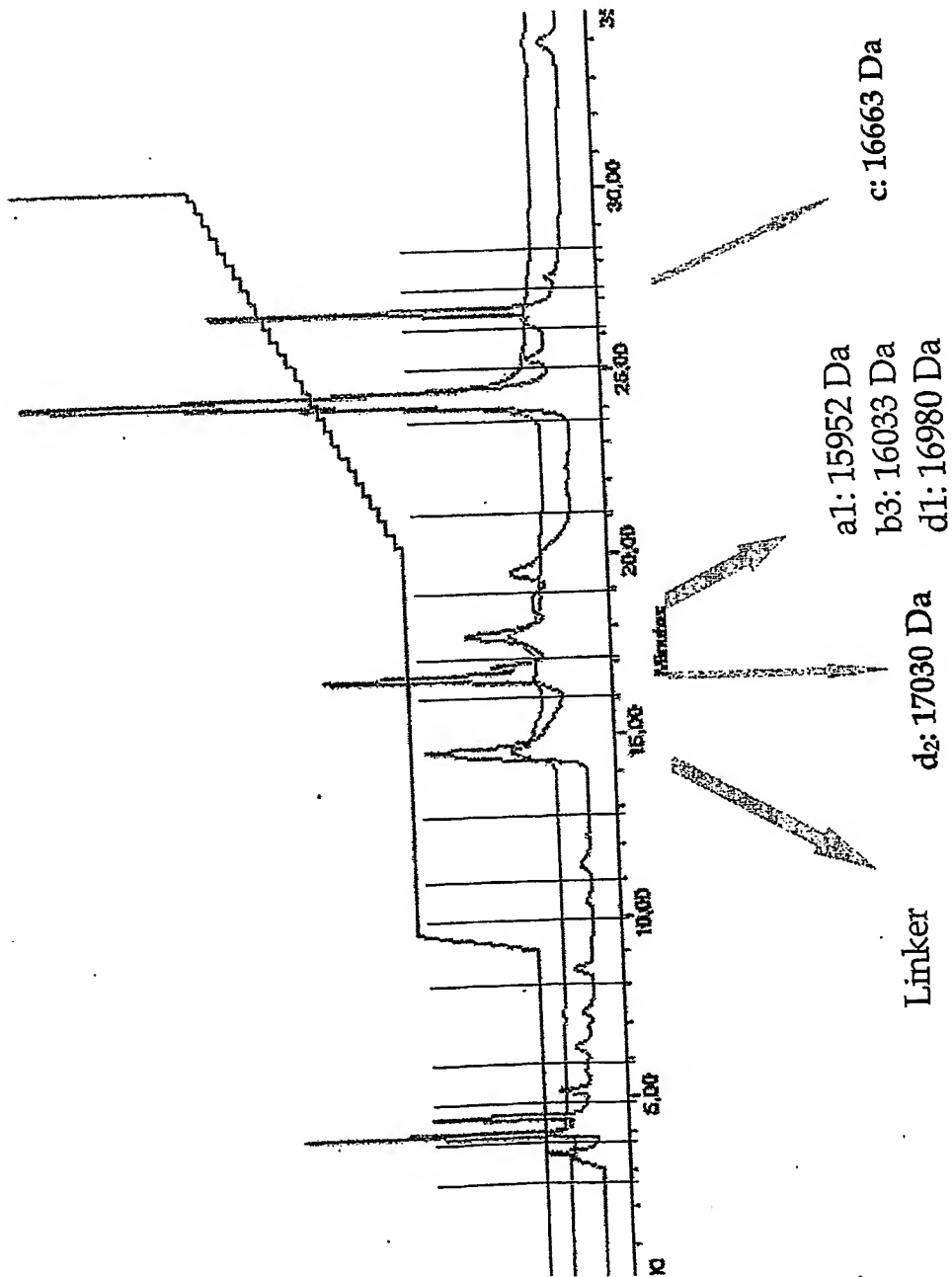


FIGURE 12

LISTE DE SEQUENCES

<110> CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

<120> GLOB

<130> IFB 03 CA CNR GLOB

<160> 18

<170> PatentIn version 3.1

<210> 1

<211> 376

<212> ADN

<213> Arenicola marina

<220>

<221> CDS

<222> (1)..(375)

<223>

<400> 1

gga cct ttg cag cgc ctc ctg gtc aag acc cag tgg aac aag gtg tac	48
Gly Pro Leu Gln Arg Leu Leu Val Lys Thr Gln Trp Asn Lys Val Tyr	
1 5 10 15	

ggc acc agc aag gtc agg gac gag gcc gga cac gtc ctc tgg aag gct	96
Gly Thr Ser Lys Val Arg Asp Glu Ala Gly His Val Leu Trp Lys Ala	
20 25 30	

att ttc gcc cag gat ccc gag acc cgg gct ctc ttc aag aga gtc aac	144
Ile Phe Ala Gln Asp Pro Glu Thr Arg Ala Leu Phe Lys Arg Val Asn	
35 40 45	

ggt gac gac atc tac tct ccc gag ttc atg gct cac agc gcc cgt gtc	192
Gly Asp Asp Ile Tyr Ser Pro Glu Phe Met Ala His Ser Ala Arg Val	
50 55 60	

ttg ggt ggc ctt gac att gcc atc tcc ctc ctc gac aac cag gct gac	240
Leu Gly Gly Leu Asp Ile Ala Ile Ser Leu Leu Asp Asn Gln Ala Asp	
65 70 75 80	

ctt gac gtc gcc ctg gct cac ctt cac gtg cag cac gta gad agg cac	288
Leu Asp Val Ala Leu Ala His Leu His Val Gln His Val Glu Arg His	
85 90 95	

atc cca acc cgc tac ttc gat ctg ttc aag aac gcc ctg atg gag tat	336
Ile Pro Thr Arg Tyr Phe Asp Leu Phe Lys Asn Ala Leu Met Glu Tyr	
100 105 110	

gcc ccc agc gcc ctg gga cgc tgc ttt gat aag gac gca t	376
Ala Pro Ser Ala Leu Gly Arg Cys Phe Asp Lys Asp Ala	
115 120 125	

<210> 2

<211> 125

<212> PRT

<213> Arenicola marina

<400> 2
 Gly Pro Leu Gln Arg Leu Leu Val Lys Thr Gln Trp Asn Lys Val Tyr
 1 5 10 15
 Gly Thr Ser Lys Val Arg Asp Glu Ala Gly His Val Leu Trp Lys Ala
 20 25 30
 Ile Phe Ala Gln Asp Pro Glu Thr Arg Ala Leu Phe Lys Arg Val Asn
 35 40 45
 Gly Asp Asp Ile Tyr Ser Pro Glu Phe Met Ala His Ser Ala Arg Val
 50 55 60
 Leu Gly Gly Leu Asp Ile Ala Ile Ser Leu Leu Asp Asn Gln Ala Asp
 65 70 75 80
 Leu Asp Val Ala Leu Ala His Leu His Val Gln His Val Glu Arg His
 85 90 95
 Ile Pro Thr Arg Tyr Phe Asp Leu Phe Lys Asn Ala Leu Met Glu Tyr
 100 105 110
 Ala Pro Ser Ala Leu Gly Arg Cys Phe Asp Lys Asp Ala
 115 120 125

<210> 3
 <211> 288
 <212> ADN
 <213> Arenicola marina
 <220>
 <221> CDS
 <222> (1) .. (288)
 <223>

<400> 3
 tgc gga ccc ctt cag cgc ctg aag gtc aag cgc cag tgg gct gag gct 48
 Cys Gly Pro Leu Gln Arg Leu Lys Val Lys Arg Gln Trp Ala Glu Ala 15
 1 5 10
 tat gga agc gga aac agc agg gag gaa ttc gga cac ttc atc tgg tcc 96
 Tyr Gly Ser Gly Asn Ser Arg Glu Glu Phe Gly His Phe Ile Trp Ser 30
 20 25
 cat gtc ttc cag cac tcg cct gct gcc cgc gac atg ttc aag cgc gtc 144
 His Val Phe Gln His Ser Pro Ala Ala Arg Asp Met Phe Lys Arg Val 45
 35 40
 cgc ggt gac aac atc cac acc cca gca ttc atg gcc cac gcc acc cgt 192
 Arg Gly Asp Asn Ile His Thr Pro Ala Phe Met Ala His Ala Thr Arg 60
 50 55
 gtg ctc ggt gga ctc gac atg tgc att gcc ctt ctc gat gat gaa ccc 240
 Val Leu Gly Gly Leu Asp Met Cys Ile Ala Leu Leu Asp Asp Glu Pro 80
 65 70 75
 gtt ctg aac acg cag ctc gct cat ctt gcc aag caa cac gaa acc cgt 288
 Val Leu Asn Thr Gln Leu Ala His Leu Ala Lys Gln His Glu Thr Arg 95
 85 90

<210> 4
 <211> 96
 <212> PRT
 <213> Arenicola marina

<400> 4
 Cys Gly Pro Leu Gln Arg Leu Lys Val Lys Arg Gln Trp Ala Glu Ala
 1 5 10 15

Tyr Gly Ser Gly Asn Ser Arg Glu Glu Phe Gly His Phe Ile Trp Ser
 20 25 30

His Val Phe Gln His Ser Pro Ala Ala Arg Asp Met Phe Lys Arg Val
 35 40 45

Arg Gly Asp Asn Ile His Thr Pro Ala Phe Met Ala His Ala Thr Arg
 50 55 60

Val Leu Gly Gly Leu Asp Met Cys Ile Ala Leu Leu Asp Asp Glu Pro
 65 70 75 80

Val Leu Asn Thr Gln Leu Ala His Leu Ala Lys Gln His Glu Thr Arg
 85 90 95

<210> 5
 <211> 360
 <212> ADN
 <213> Arenicola marina

<220>
 <221> CDS
 <222> (1)..(360)
 <223>

<400> 5
 aag gtg aag cac caa tgg gtg cag gtg tac agc ggc cat ggt tac gag 48
 Lys Val Lys His Gln Trp Val Gln Val Tyr Ser Gly His Gly Tyr Glu
 1 5 10 15
 cgt gag gcg ttc ggc aga gag gtc ttc ctc gag atg tac aac cag gca 96
 Arg Glu Ala Phe Gly Arg Glu Val Phe Leu Glu Met Tyr Asn Gln Ala
 20 25 30
 ccc aag gcc aag gac ctc ttc acc agg gtc agg ggc gag aac gtc ttc 144
 Pro Lys Ala Lys Asp Leu Phe Thr Arg Val Arg Gly Glu Asn Val Phe
 35 40 45
 tcc ccc gag ttc gga gcc cac atg gtc cgt gtg ctc gga ggt ctc gac 192
 Ser Pro Glu Phe Gly Ala His Met Val Arg Val Leu Gly Gly Leu Asp
 50 55 60
 atg tgc atc gct ctg ctg tcc gat gac acc gtc ctc aac gcc cag ctt 240
 Met Cys Ile Ala Leu Leu Ser Asp Asp Thr Val Leu Asn Ala Gln Leu
 65 70 75 80
 gct cac ctc agc acg cag cac aag gac cgt gga atc ccc aac gag tac 288
 Ala His Leu Ser Thr Gln His Lys Asp Arg Gly Ile Pro Asn Glu Tyr
 85 90 95

ttc gat gtg gtg aag gtc gcc ctc atg aag gtc gtc ccc ggc cac gtt 336
 Phe Asp Val Val Lys Val Ala Leu Met Lys Val Val Pro Gly His Val
 100 105 110

tca cac ttc gat atc ggc gcg tgg 360
 Ser His Phe Asp Ile Gly Ala Trp
 115 120

<210> 6
 <211> 120
 <212> PRT
 <213> Arenicola marina

<400> 6
 Lys Val Lys His Gln Trp Val Gln Val Tyr Ser Gly His Gly Tyr Glu
 1 5 10 15

Arg Glu Ala Phe Gly Arg Glu Val Phe Leu Glu Met Tyr Asn Gln Ala
 20 25 30

Pro Lys Ala Lys Asp Leu Phe Thr Arg Val Arg Gly Glu Asn Val Phe
 35 40 45

Ser Pro Glu Phe Gly Ala His Met Val Arg Val Leu Gly Gly Leu Asp
 50 55 60

Met Cys Ile Ala Leu Leu Ser Asp Asp Thr Val Leu Asn Ala Gln Leu
 65 70 75 80

Ala His Leu Ser Thr Gln His Lys Asp Arg Gly Ile Pro Asn Glu Tyr
 85 90 95

Phe Asp Val Val Lys Val Ala Leu Met Lys Val Val Pro Gly His Val
 100 105 110

Ser His Phe Asp Ile Gly Ala Trp
 115 120

<210> 7
 <211> 390
 <212> ADN
 <213> Arenicola marina

<220>
 <221> CDS
 <222> (1)..(390)
 <223>

<400> 7
 tgt tgc agt ata gag gac cgc aag gag gtc cag acg ctg tgg agt gag 48
 Cys Cys Ser Ile Glu Asp Arg Lys Glu Val Gln Thr Leu Trp Ser Glu
 1 5 10 15

atc tgg agt gcc cag ttc act ggt cgc cgt gtc cag gtt gcc cag gct 96
 Ile Trp Ser Ala Gln Phe Thr Gly Arg Arg Val Gln Val Ala Gln Ala
 20 25 30

gtg ttc gag gac ctc ttc cgc cgc gac ccc gag tcc aag aac ctg ttc 144
 Val Phe Glu Asp Leu Phe Arg Arg Asp Pro Glu Ser Lys Asn Leu Phe
 35 40 45

aag cgc gtc aat gtt gac gac atg aac agc ccc gaa ttc cac gct cac 192
 Lys Arg Val Asn Val Asp Asp Met Asn Ser Pro Glu Phe His Ala His
 50 55 60

tgc atc cgt gtt gtc aac ggt ctt gac acc gtg atc ggt ctc ctt gac 240
 Cys Ile Arg Val Val Asn Gly Leu Asp Thr Val Ile Gly Leu Leu Asp
 65 70 75 80

gac ccc gac acc ctg aag tcc cag ctc gag cac ttg gcc cag cag cac 288
 Asp Pro Asp Thr Leu Lys Ser Gln Leu Glu His Leu Ala Gln Gln His
 85 90 95

aag gag cgt gat ggc atc cac aag acc cac ttc gac gag atg tcc cac 336
 Lys Glu Arg Asp Gly Ile His Lys Thr His Phe Asp Glu Met Ser His
 100 105 110

gcc ttc ggc gcc gtc atg ccc cag gtc agc agc tgt ttc aac ccc gat 384
 Ala Phe Gly Ala Val Met Pro Gln Val Ser Ser Cys Phe Asn Pro Asp
 115 120 125

gca tga 390
 Ala

<210> 8
 <211> 129
 <212> PRT
 <213> Arenicola marina

<400> 8
 Cys Cys Ser Ile Glu Asp Arg Lys Glu Val Gln Thr Leu Trp Ser Glu
 1 5 10 15

Ile Trp Ser Ala Gln Phe Thr Gly Arg Arg Val Gln Val Ala Gln Ala
 20 25 30

Val Phe Glu Asp Leu Phe Arg Arg Asp Pro Glu Ser Lys Asn Leu Phe
 35 40 45

Lys Arg Val Asn Val Asp Asp Met Asn Ser Pro Glu Phe His Ala His
 50 55 60

Cys Ile Arg Val Val Asn Gly Leu Asp Thr Val Ile Gly Leu Leu Asp
 65 70 75 80

Asp Pro Asp Thr Leu Lys Ser Gln Leu Glu His Leu Ala Gln Gln His
 85 90 95

Lys Glu Arg Asp Gly Ile His Lys Thr His Phe Asp Glu Met Ser His
 100 105 110

Ala Phe Gly Ala Val Met Pro Gln Val Ser Ser Cys Phe Asn Pro Asp
 115 120 125

Ala

<210> 9
 <211> 20
 <212> ADN
 <213> séquence artificielle

<220>
 <223> amorce PCR

<220>
 <221> misc_feature
 <222> (3)..(3)
 <223> A ou G

<220>
 <221> misc_feature
 <222> (6)..(6)
 <223> C ou T

<220>
 <221> misc_feature
 <222> (9)..(9)
 <223> A, G, C ou T

<220>
 <221> misc_feature
 <222> (12)..(12)
 <223> A, G, C ou T

<220>
 <221> misc_feature
 <222> (15)..(15)
 <223>

<220>
 <221> misc_feature
 <222> (15)..(15)
 <223> A ou G

<400> 9
 gartgyggnc cnttrcarcg

20

<210> 10
 <211> 21
 <212> ADN
 <213> séquence artificielle

<220>
 <223> amorce PCR

<220>
 <221> misc_feature
 <222> (4)..(4)
 <223> A, G, C ou T

<220>
 <221> misc_feature

<222> (7)..(7)
<223> A, G, C ou T

<220>
<221> misc_feature
<222> (10)..(10)
<223> C ou T

<220>
<221> misc_feature
<222> (13)..(13)
<223> A ou G

<220>
<221> misc_feature
<222> (16)..(16)
<223> A ou G

<400> 10
ccangntcy ttrtcraagc a

21

<210> 11
<211> 19
<212> ADN
<213> séquence artificielle

<220>
<223> amorce PCR

<220>
<221> misc_feature
<222> (2)..(2)
<223> A, G, C ou T

<220>
<221> misc_feature
<222> (5)..(5)
<223> C ou T

<220>
<221> misc_feature
<222> (8)..(8)
<223> A, G, C ou T

<220>
<221> misc_feature
<222> (11)..(11)
<223> A, G, C ou T

<220>
<221> misc_feature
<222> (14)..(14)
<223> A, G, C ou T

<220>
<221> misc_feature
<222> (17)..(17)
<223> A ou G

<400> 11
antgyggncc nctncarcg

19

<210> 12
<211> 21
<212> ADN
<213> séquence artificielle

<220>
<223> amorce PCR

<220>
<221> misc_feature
<222> (4)..(4)
<223> A, G, C ou T

<220>
<221> misc_feature
<222> (7)..(7)
<223> A, G, C ou T

<220>
<221> misc_feature
<222> (10)..(10)
<223> C ou T

<220>
<221> misc_feature
<222> (13)..(13)
<223> A ou G

<220>
<221> misc_feature
<222> (16)..(16)
<223> A ou G

<400> 12
ccangcntcy ttrtcraagc a

21

<210> 13
<211> 18
<212> ADN
<213> séquence artificielle

<220>
<223> amorce PCR

<220>
<221> misc_feature
<222> (3)..(3)
<223> A ou G

<220>
<221> misc_feature
<222> (6)..(6)
<223> inosine

<220>
<221> misc_feature

<222> (9)..(9)

<223> A ou G

<220>

<221> misc_feature

<222> (12)..(12)

<223> A, G, C ou T

<400> 13

aargtiaarc anaactgg

18-

<210> 14

<211> 18

<212> ADN

<213> séquence artificielle

<220>

<223> amorce PCR

<220>

<221> misc_feature

<222> (4)..(4)

<223> A, G, C ou T

<220>

<221> misc_feature

<222> (7)..(7)

<223> A, G, C ou T

<220>

<221> misc_feature

<222> (10)..(10)

<223> A, G ou T

<220>

<221> misc_feature

<222> (13)..(13)

<223> A ou G

<220>

<221> misc_feature

<222> (16)..(16)

<223> A ou G

<400> 14

ccangcnccd atrtcraa

18

<210> 15

<211> 20

<212> ADN

<213> séquence artificielle

<220>

<223> amorce PCR

<220>

<221> misc_feature

<222> (3)..(3)

<223> C ou T

<220>
 <221> misc_feature
 <222> (6)..(6)
 <223> C ou T

<220>
 <221> misc_feature
 <222> (9)..(9)
 <223> C ou T

<220>
 <221> misc_feature
 <222> (18)..(18)
 <223> C ou T

<220>
 <221> misc_feature
 <222> (12)..(12)
 <223> A, C ou T

<220>
 <221> misc_feature
 <222> (15)..(15)
 <223> A ou G

<400> 15
 tgytgyagya thgargaycg

20

<210> 16
 <211> 20
 <212> ADN
 <213> séquence artificielle

<220>
 <223> amorce PCR

<220>
 <221> misc_feature
 <222> (3)..(3)
 <223> A, G, C ou T

<220>
 <221> misc_feature
 <222> (6)..(6)
 <223> A, G, C ou T

<220>
 <221> misc_feature
 <222> (7)..(7)
 <223> C ou T

<220>
 <221> misc_feature
 <222> (9)..(9)
 <223> A ou G

<220>
 <221> misc_feature

<222> (12)..(12)
<223> A ou G

<220>
<221> misc_feature
<222> (15)..(15)
<223> A ou G

<220>
<221> misc_feature
<222> (18)..(18)
<223> A ou G

<400> 16
cangcnyrc trttraarca

20

<210> 17
<211> 17
<212> ADN
<213> séquence artificielle

<220>
<223> amorce PCR

<220>
<221> misc_feature
<222> (3)..(3)
<223> C ou T

<220>
<221> misc_feature
<222> (6)..(6)
<223> C ou T

<220>
<221> misc_feature
<222> (9)..(9)
<223> C ou T

<220>
<221> misc_feature
<222> (12)..(12)
<223> A, G, C ou T

<220>
<221> misc_feature
<222> (15)..(15)
<223> C ou T

<400> 17
gaygaytgyg tntgycc

17

<210> 18
<211> 19
<212> DNA
<213> séquence artificielle

<220>
<223> amorce PCR

<220>
 <221> misc_feature
 <222> (2)..(2)
 <223> A, G, C ou T

<220>
 <221> misc_feature
 <222> (14)..(14)
 <223> A, G, C ou T

<220>
 <221> misc_feature
 <222> (5)..(5)
 <223> A ou G

<220>
 <221> misc_feature
 <222> (8)..(8)
 <223> A ou G

<220>
 <221> misc_feature
 <222> (11)..(11)
 <223> A ou G

<220>
 <221> misc_feature
 <222> (17)..(17)
 <223> A ou G

<400> 18
 anacraarte rtcncrca

DÉPARTEMENT DES BREVETS

26 bis, rue de Saint Pétersbourg

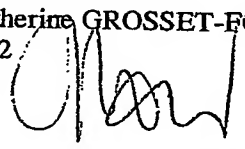
75800 Paris Cedex 08

Téléphone : 01 53 04 53 04 Télécopie : 01 42 94 86 54

DÉSIGNATION D'INVENTEUR(S) Page N° 1/2

(Si le demandeur n'est pas l'inventeur ou l'unique inventeur)

Cet imprimé est à remplir lisiblement à l'encre noire. 08 113 W / 260899

Vos références pour ce dossier (facultatif)		IFB 03 CA CNR GLOB	
N° D'ENREGISTREMENT NATIONAL		03/11992	
TITRE DE L'INVENTION (200 caractères ou espaces maximum)			
<p>PROCÉDE DE DISSOCIATION DE LA MOLECULE D'HEMOGLOBINE EXTRACELLULAIRE D'ARENICOLA MARINA, CHAÎNES PROTEIQUES CONSTITUANT LADITE MOLECULE ET SEQUENCES NUCLEOTIDIQUES CODANT POUR LESDITES CHAÎNES PROTEIQUES</p>			
LE(S) DEMANDEUR(S) :			
<p>CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE 3, rue Michel-Ange F-75794 PARIS CEDEX 16, France</p>			
DESIGNE(NT) EN TANT QU'INVENTEUR(S) : (Indiquez en haut à droite «Page N° 1/1» S'il y a plus de trois inventeurs, utilisez un formulaire identique et numérotez chaque page en indiquant le nombre total de pages).			
Nom		ZAL	
Prénoms		Franck	
Adresse	Rue	Saint Kirio	
	Code postal et ville	26900 MORLAIX - PLOUJEAN	
Société d'appartenance (facultatif)			
Nom		CHABASSE	
Prénoms		Christine	
Adresse	Rue	16, rue Gambetta	
	Code postal et ville	29680 ROSCOFF	
Société d'appartenance (facultatif)			
Nom		ROUSSELOT	
Prénoms		Morgane	
Adresse	Rue	2, rue Dociatis	
	Code postal et ville	29250 SAINT POL DE LEON	
Société d'appartenance (facultatif)			
DATE ET SIGNATURE(S) DU (DES) DEMANDEUR(S) DU DU MANDATAIRE Nom et qualité du signataire)		<p>Paris, le 27 Novembre 2003</p> <p>Chantal, Catherine GROSSET-FOURNIER, Mandataire 422.5/PP112</p> 	

TEMENT DES BREVETS

rue de Saint Pétersbourg
Paris Cedex 08
télé : 01 53 04 53 04 Télécopie : 01 42 94 86 54

DÉSIGNATION D'INVENTEUR(S) Page N° 2./2
(Si le demandeur n'est pas l'inventeur ou l'unique inventeur)

Cet imprimé est à remplir lisiblement à l'encre noire

08 113 W / 260899

références pour ce dossier (facultatif)	IFB 03 CA CNR GLOB
D'ENREGISTREMENT NATIONAL	03/11992
RE DE L'INVENTION (200 caractères ou espaces maximum)	

PROCÉDE DE DISSOCIATION DE LA MOLECULE D'HEMOGLOBINE EXTRACELLULAIRE
D'ARENICOLA MARINA, CHAÎNES PROTEIQUES CONSTITUANT LADITE MOLECULE ET SEQUENCES
NUCLEOTIDIQUES CODANT POUR LESDITES CHAÎNES PROTEIQUES

(S) DEMANDEUR(S) :

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE
3, rue Michel-Ange
F-75794 PARIS CEDEX 16, France


SIGNE(NT) EN TANT QU'INVENTEUR(S) : (Indiquez en haut à droite «Page N° 1/1» S'il y a plus de trois inventeurs,
lisez un formulaire identique et numérotez chaque page en indiquant le nombre total de pages).

Nom		BAILLY	
Prénoms		Xavier	
Adresse	Rue	71, rue du Poulguegen	
	Code postal et ville	29250 SANTEC	
Société d'appartenance (facultatif)			
Nom			
Prénoms			
Adresse	Rue		
	Code postal et ville		
Société d'appartenance (facultatif)			
Nom			
Prénoms			
Adresse	Rue		
	Code postal et ville		
Société d'appartenance (facultatif)			

DATE ET SIGNATURE(S)
DU (DES) DEMANDEUR(S)
OU DU MANDATAIRE
(Nom et qualité du signataire)

Paris, le 27 novembre 2003

Chantal GROSSET-FOURNIER, Mandataire
422.5/PP112



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☒ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.